ajh

DEPARTMENT OF HEALTH AND HUMAN SERVICES

PUBLIC HEALTH SERVICE

FOOD AND DRUG ADMINISTRATION

ANTI-INFECTIVE DRUGS ADVISORY COMMITTEE MEETING

64TH MEETING

ISSUE:

GUIDANCE DOCUMENTS ON DEVELOPING ANTIMICROBIAL DRUGS:

GENERAL CONSIDERATIONS & INDIVIDUAL INDICATIONS

Wednesday, July 29, 1998

8:10 a.m.

Hilton Hotel
Grand Ballroom

MILLER REPORTING COMPANY, INC.
507 C Street, N.E.
Washington, D.C. 20002
(202) 546-6666

620 Perry Parkway
Gaithersburg, Maryland

PARTICIPANTS

William A. Craig, M.D., Chairperson
Ermona McGoodwin, Executive Secretary

MEMBERS

Patricia J. Chesney, M.D.
Celia D. Christie, M.D.
Nancy K. Henry, M.D., Ph.D.
Barbara E. Murray, M.D.
Carl W. Norden, M.D.
Keith A. Rodvold, Pharm. D. (Consumer Rep)
David E. Soper, M.D.

CONSULTANTS

William Blackwelder, Ph.D.
Barth Reller, M.D.
Janet T. Wittes, Ph.D.

FDA

Renata Albrecht, M.D.
Gary Chikami, M.D. (a.m. only)
Dan Davis, M.D. (p.m. only)
Brad Leissa, M.D. (p.m. only)
Mark Goldberger, M.D., Ph.D.
Dianne Murphy, M.D. (a.m. only)
Joseph Winfield, M.D. (p.m. only)

C O N T E N T S

P R O C E E D I N G S

Call to Order

DR. CRAIG: Good morning. I would like to welcome you to the Anti-infective Drugs Advisory Committee Meeting, the 64th one. To begin the program, what I want to do is have everybody at the table get their name onto the record. We will start with Dr. Albrecht.

DR. ALBRECHT: Good morning. I am Renata Albrecht.

DR. CHIKAMI: I am Gary Chikami. I am the Director of the Division of Anti-infective Drug Products.

DR. MURPHY: I am Dianne Murphy. I am Office Director for ODE IV.

DR. GOLDBERGER: Mark Goldberger, Director of the Division of Special Pathogens.

DR. MURRAY: Barbara Murray. I am on the Committee. I'm from the University of Texas Medical School in Houston.

DR. RELLER: Barth Reller, Duke University Medical Center.

MS. McGOODWIN: Ermona McGoodwin, FDA.

DR. CRAIG: Bill Craig, University of Wisconsin.

DR. NORDEN: Carl Norden, Cooper Hospital/ University of New Jersey Medical Center.

DR. CHRISTIE: Good morning. I am Celia Christie. I am from the University of Cincinnati and the Children's Hospital Medical Center in Cincinnati.

DR. HENRY: Nancy Henry, Department of Pediatrics, Mayo Clinic.

DR. RODVOLD: Keith Rodvold, University of Illinois at Chicago.

DR. SOPER: David Soper, Medical University, South Carolina.

DR. CHESNEY: Joan Chesney, the University of Tennessee in Memphis.

DR. WITTES: Janet Wittes, Statistics Collaborative in D.C.

DR. BLACKWELDER: Bill Blackwelder from NIH.

DR. CRAIG: Thank you very much.

Ermona McGoodwin will then read the conflict of interest statement.

Conflict of Interest Statement

MS. McGOODWIN: Thank you, Dr. Craig. The following announcement addresses the issue of conflict of interest with regard to this meeting and is made a part of the record to preclude even the appearance of such at this meeting.

In accordance with 18 USC 208, general matters

waivers have been granted to all committee participants who have interests in companies or organizations which could be affected by the committee's decisions of guidance documents for guidance to industry on antimicrobial drug products for the treatment of infections.

A copy of these waiver statements may be obtained by submitting a written request to the agency's Freedom of Information Office, Room 12A-30, Parklawn Building.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants are aware of the need to exclude themselves from such involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask, in the interest of fairness, that they address any current or previous financial involvement with any firm whose product they may wish to comment upon.

Thank you.

DR. CRAIG: Thank you, Ermona.

The welcome and introduction will be given by Dianne Murphy.

Welcome and Introduction

DR. MURPHY: I wanted to welcome the committee,

the consultants, the guests, and the audience to what is truly a massive effort. Today, we are going to be involved, and the next two days, in evolving, enhancing, and developing guidance for the development of antimicrobial drug products.

You will hear from Dr. Albrecht how we got to this point, a little bit of history to put it in perspective. Then, I will give you a general overview of what we hope to achieve during this session in the next two days, and then we will get down to the specifics with Dr. Chikami and Dr. Lin presenting some of the foundations, if you will, of how we plan to reach our goals.

During the next two days, we will be providing you specific indications to review and to provide your comments to us and your advice and guidance. At the end of this session, we will have a 90-day comment period, and the FDA will then review those comments and once again publish these guidances.

I am not going to say too much more because we have a lot of work ahead of us.

I would like Dr. Albrecht to please get us started with an overview of where we have been and where we are going to go, and also to tell you all that she is really responsible for coordinating, producing, and getting this

all together, and will be our guide over the next couple of days.

Thank you.

DR. CRAIG: Thank you, Dr. Murphy.

Dr. Albrecht.

Guidance Development

DR. ALBRECHT: Thank you, Dr. Craig. Thank you, Dr. Murphy, for the introduction.

[Slide.]

Good morning, Dr. Craig, members of the Committee, consultants and colleagues from industry and FDA. The purpose of my presentation this morning is to give you an overview of the guidance to industry - developing antimicrobial drugs process from the FDA.

My name is Renata Albrecht. I am the Deputy Director in the Division of Special Pathogens and Immunologic Drug Products. As Dr. Murphy mentioned, I have coordinated and led the guidance to industry - developing antimicrobial drugs effort for approximately the last year and a half.

So, as a colleague fondly said, that means, "If anything goes wrong, you are the one we blame."

[Slide.]

In the next three days, you are going to hear

approximately 20 presentations by FDA colleagues. Many of these, approximately a dozen, will be followed by comments by members of the advisory committee or by invited consultants. We will have discussions of specific topics during these sessions and, after each, there will be the opportunity for questions and comments.

In addition, on each day, we have reserved approximately a half an hour for open public hearing in case questions and spontaneous issues do come up.

[Slide.]

Before going into the background, let me mention that, as you look at the agendas, you will notice that we have very long days today and tomorrow but, on Friday, we are going to try to finish earlier in the day to enable everyone who has taken all this time out of their busy schedule to return back home.

By way of background; the FDA, and specifically the divisions within the Office of Drug Evaluation IV, have a fairly long history of interacting with industry and others in the development of drugs and providing advice in the forms of letters, meetings, both at the FDA and advisory committee meetings as we had last March, and written guidelines, which is what we termed them in the past, written guidelines and points to consider to try to assist

companies in developing antimicrobial drugs.

The initial document that was provided in written form as a guideline was in 1977. It is called Clinical Evaluation of Anti-Infective Drugs Systemic. In 1992, in a joint FDA/IDSA effort, the 1992 IDSA/FDA guidelines were written and published.

At the same time, or within an month of that effort coming to publication, the Division of Anti-Infective Drug Products, then under the leadership of Dr. Mack Lumpkin, published the Points to Consider document. That document was also discussed in an advisory committee proceeding, the issues of the Points to Consider document.

And then, in 1997, the Division of Anti-Infective Drug Products published the guidance document on evaluating clinical trials. The content of that document was presented at the advisory committee of March 1997.

I mention this, that the effort did start in the Division of Anti-Infective Drugs, because it will be evident as we go through the specific indications over the next three days that there is a focus on antibacterial infections and that is of historical interest.

[Slide.]

Although the effort did, in the last few years, involve the Division of Anti-Infective Drugs, it has now

been expanded and, under the leadership of Dr. Dianne Murphy, it is now an effort that involves both the Divisions of Anti-Infective Drug Products, the Division of Special Pathogens and Immunologic Drug Products and the Division of Antiviral Drug Products.

[Slide.]

What is the current environment? We continue to provide information to industry on developing antimicrobial drugs, and most of you aware that on November 21, 1997, the FDA Modernization Act was passed.

It has some provisions that are relevant and have modified some of the approaches that we use, and Dr. Murphy will actually highlight some of the sections that are applicable to our three divisions.

In addition, shortly after the issuance of the 1997 guidance by the Division of Anti-Infective Drugs, the agency published another Federal Register notice on Good Guidance Practices, and we are operating in the context of those guidances.

There are other relevant and applicable documents provided by the agency including the Clinical Effectiveness Guidance, which was published this year, and of course there are many ICH efforts that relate to the activities that we are undertaking.

[Slide.]

The most directly applicable Federal Register notice to the process that we are discussing in the next three days is the Federal Register notice that published on July 21st of this year, and copies have been made available to the committee members for their perusal.

The Federal Register notice announces an ODE IV effort to revise and update existing guidance documents. As part of that process, it is the intent of ODE IV to create a series of specific guidance documents and a general overview document, and in fact, as of today, the public notice, the Federal Register notice issued announcing the availability of 18 specific guidance documents.

Most of you are of course already aware of those because they were posted on the FDA home page last week. As part of the GGP effort, it is the intent of the agency to always discuss in public these documents for public comment, and the ODE IV effort is being presented at this advisory committee.

Also, in the Federal Register notice, request is made that interested parties comment both on the proposal of writing these guidance documents, as well as on the actual content of the individual documents.

[Slide.]

For those who remember the guidance document we published last year, it was a large tome of approximately 100 pages with many individual sections within it. We believe it is more practical to have reformatted that into individual documents, a larger overview document called the General Considerations for Clinical Trials Guidance, and a series of smaller ones.

In the large document, we covered general areas in topics of microbiology, pharmacology/toxicology, chemistry, clinical pharmacology, clinical issues such as protocol design and analyses, and biostatistical issues.

Over the next three days, you will hear FDA staff discuss the highlights of these individual subsections of the large document.

[Slide.]

In addition to the General Considerations guidance document, we have 17 so-called companion documents developing antimicrobials for the treatment of a variety of infections. Many of these should look fairly familiar because they cover topics that were presented at the 1997 Anti-Infective Advisory Committee. Others are going to be presented for the first time during the next three days.

[Slide.]

Many of you have already discovered these

documents are available on the FDA home page. The address for the documents is given at the top. It was very kind of the webmaster to actually create a specific site to post the 18 documents in one place.

[Slide.]

This is a busy slide and I did this almost intentionally because, as Dr. Murphy said, this has been really a daunting task and an extraordinary effort by many, many people, and I don't know if we have set a record from the office and the three divisions on how many documents we have posted, however, there are many, and we hope that they will be received as they were intended, as guidance.

The topics from the previous advisory committee and now put out in sort of a final draft version are the acute otitis document, acute sinusitis, acute exacerbation of chronic bronchitis, secondary bacterial infections, acute bronchitis, community-acquired pneumonia and nosocomial pneumonia, uncomplicated gonorrhea, uncomplicated urinary tract infection, uncomplicated and complicated skin and skin structure infections, and to be presented at the advisory committee for the first time are the following documents: meningitis, vulvovaginal candidiasis, bacterial vaginosis, streptococcal pharyngitis and tonsillitis, complicated UTI and pyelonephritis, bacterial prostatitis, early Lyme

disease, and the empiric therapy of febrile neutropenia.

[Slide.]

This effort of course has been made possible through the hard work of many, many people. I almost hesitate to start these things because I always leave out the critical people, but in fact most of staff in the Division of Anti-Infective Drug Products and the Division of Special Pathogens has been involved in this effort, staff from the Office of Policy, advisers and consultants, and, of course, our advisory committee consultants.

I would like to mention three individuals whose hard work over the last months, if not years, has facilitated my task, and they are Dr. Lillian Gavrilovich, Deputy Director of the Division of Anti-Infective Drugs, who actually was asked to launch this effort back in 1996; Dr. Juanita Fastman, who made sure that the weekly meetings we have had over the last year or so, we had a conference room available; and finally, Nancy Derr, in the Office of Policy, without whom you would not have those well organized, well formatted documents for reading.

[Slide.]

A few comments about guidance, what is a guidance. We usually say that guidance represents our current thinking on a particular topic or topics. A guidance is not a law,

it is not a regulation, it does not have the force of law behind it per se. It is not considered legally binding, it is considered a device.

The documents that are posted now are in draft form, which means that there is still time for revisions and comments. We are using this form of the advisory committee meeting to present them publicly, and there is a 90-day comment period where we invite public written comment on the content of the documents.

It is the intent of ODE IV to then review these comments and revise, as well as finalize, the existing guidance documents.

[Slide.]

The content of the documents has been revised and information has been incorporated from a variety of sources including previous guideline documents, recent experiences, the published literature, and advice from consultants and industry.

[Slide.]

Let me mention briefly the format of the guidance. We do have style guides in the FDA for guidance documents, and then also within our antimicrobial drug development guidances we wanted to organize the information in such a way that it was easy to retrieve and use.

So, you will notice that the individual indication companion documents have a regulatory background section, a study considerations section, information on specific inclusion and exclusion criteria subsections, sections about drugs and dosage regimens, information on evaluation visits, outcome, and statistics.

[Slide.]

What about the next three days? We are going to be hearing FDA presentations, comments by committee members and consultants. Clearly, it is not possible to cover everything. So, our plan is that on topics that have been previously presented, we will provide updates, basically, a summary of comments from the last advisory committee, and comments that were submitted to us from industry in response to the last guidance document.

For new topics, indications that have not been previously presented, we will give more complete presentations. Individual presentations will in many cases be followed by discussion from the committee and consultants. There will then be some particular issues or questions raised, and probably we should have time for audience comments, as well.

[Slide.]

Let me mention up-front that there is no plan to

present the following topics: chemistry, uncomplicated urinary tract infections, uncomplicated gonorrhea, and uncomplicated and complicated skin and skin structure infections. These documents received either no or basic comments that were taken into consideration. These were incorporated into the documents, and they are now presented for final comment.

[Slide.]

The topics that we will present, when we organized this agenda, we tackled how to organize it, what should we do first and what should we do next, and the thought was that those topics that have not previously been discussed, we would try to put earlier in the course of the meeting, so that, for example, today, you will be hearing about meningitis, vulvovaginal candidiasis, and bacterial vaginosis.

In addition, there was an attempt to cluster these categories, so that if people have a particular interest in an area, they might be able to plan their schedules accordingly, so, for example, we have a cluster of gynecological infections, genitourinary and respiratory, and some of it was guided by the availability of our committee and consultants as to what days they could take out of their busy schedules to come and join us.

[Slide.]

Let me also mention to you as you have looked over the guidance documents, you noticed that they are really quite uniform in format, and they follow a style guide that the agency, and specifically CDER, has put forth for guidance documents. So, to try to put a little bit of lively change into the process, the presentations are very much individualized. People have used different templates and different fonts to give some different visual cues.

[Slide.]

As you noted, we do have a little bit of a change in the agenda, and I did not have an opportunity to update this slide, but after my presentation, Dr. Murphy, as she said, will discuss some basic issues on FDAMA and clinical trial design issues.

Then, Dr. Chikami will talk about clinical trials. Then, Dr. Daphne Lin will present the FDA discussion of biostatistics, and Dr. Janet Wittes will serve as consultant on that topic.

This will be followed toward the end of this morning by a presentation by Dr. Alex Rakowsky of the FDA on meningitis, and this topic will be discussed by Dr. Joan Chesney.

[Slide.]

This afternoon, we will have Dr. Brad Leissa discussing and updating us on the topic of acute otitis media, and Dr. Chesney will make further comments.

[Slide.]

Then, we will conclude this afternoon with presentations on vulvovaginal candidiasis by Dr. Joseph Winfield, of FDA, and bacterial vaginosis will be presented by Dr. Daniel Davis, of FDA. Dr. David Soper, of the committee, will serve as consultant for both of these topics.

[Slide.]

On Thursday, actually, we will start the morning--we had another change--we still start the morning with a presentation of complicated urinary tract infection and pyelonephritis by Dr. Regina Alivisatos, and that will be discussed by Dr. Barth Reller. Then, we will go through general clinical considerations. Then, Dr. Regina Alivisatos will return for a presentation on bacterial prostatitis, which will be discussed by Dr. Craig.

[Slide.]

On Thursday, this will be followed by a presentation of streptococcal pharyngitis and tonsillitis by Dr. Mamodikoe Makhene and Dr. Celia Maxwell will serve as discussant.

[Slide.]

Friday afternoon, Dr. Janice Soreth and Dr. Sousan Altaie of the FDA will present early Lyme disease, and Dr. Ray Dattwyler will discuss this topic.

That will be followed by a presentation on acute sinusitis by Dr. Eric Mann.

[Slide.]

Thursday afternoon, bronchitis, an overview will be presented by David Bostwick of FDA. Pneumonia will be presented by Dr. Alma Davidson, and a discuss of sputum gram stain will be presented by Dr. Sousan Altaie. Dr. Craig will serve as discussant on these topics.

[Slide.]

Finally, on Friday morning, we will conclude with presentations on toxicology update by Dr. Robert Osterberg, microbiology update by Dr. Sousan Altaie, and a clinical pharmacology discussion by Dr. Phil Colangelo.

[Slide.]

Last, but definitely not least, on our agenda is a presentation of empiric therapy of febrile neutropenia by Dr. David Ross, of FDA, to be discussed by Dr. Arthur Brown, consultants.

[Slide.]

At the end of that, hopefully, we will all be able

to sail on home.

With that, I would like to introduce Dr. Dianne Murphy, Office Director, ODE IV.

DR. CRAIG: Thank you, Dr. Albrecht.

Introduction and FDAMA Summary

DR. MURPHY: Once again, I welcome you all to, as you can see, a massive effort, and we really do appreciate the commitment of the committee to be with us over the three days.

The FDA reviewers have developed an intensive and hopefully challenging next three days for you and us. This meeting is a public discussion of FDA's proposed approach to drug development for antimicrobials.

You will hear presentations both on the General Considerations document tomorrow and 17 specific indications. I think you will understand probably somewhere around Friday afternoon why some of our staff affectionately refer to the series of 18 guidances as "the 18-wheeler." Sometimes they feel like they have been hit by one.

While I am on the truck theme, let's go to the first slide.

[Slide.]

This is a cartoon depicting what FDA would say were misperceptions about the review and approval process

for marketing of new therapies. On this slide, a large 18-wheeler disgorges data via paper onto the black box of the FDA where all sorts of analysis and design activities occur, and after some prolonged period of time, a decision is finally reached. So, we finally come to some decision, at least, as I said, that would be the perception of some.

This is simply not how the process works, and as you have heard, over the next three days, you will be part of the activity involved in facilitating the drug development process for antimicrobial therapy. This is more how we think it occurs, and I will go through that in just a minute.

[Slide.]

Congress, via the FDA Modernization Act, has mandated many changes including review periods for new drug applications of 6 months to 12 months. To review the large databases within these time frames, it necessitates clear communication of expectations, here, well-designed trials and efficient modes of analysis, here, and the process involves public input in many forms, including advisory committees, special public meetings, such as occurred yesterday, to seek industry's input, and publications, as you have heard, for comment of our guidances and intentions.

Companies should be consulting FDA at all stages

and particularly early in their design and trial development stages if they and we are to use our resources effectively. Good trial design is critical to this process, and you will hear more about this today.

[Slide.]

First, let me capture in three slides, the essence of 85 pages of small print concerning the FDA Modernization Act. The Act was signed into effect on November 21, 1997, not that long ago, and has a number of important sections which I will mercifully not put forth to you today, nor discuss, but I think can be summarized by the third item on here, which indicates that many of the activities of our requirements in the Act really codify ongoing FDA initiatives and existing programs.

[Slide.]

Some of the Act requirements are of particular interest to us today, such as the requirement for review times. As you can see, our NDA applications now will progressively decrease in the amount of review time from 12 months to 10 months. For serious and life-threatening diseases, we will continue with the 6-month priority review. We have been doing that for a while, and manufacturing supplements will decrease from 6 months to 4 months.

The importance of the meetings is mentioned in the

Act to enhance communication and thus, the quality and efficiency of trial design, and the development of guidance documents.

[Slide.]

Guidances are to be developed with public participation. This participation is particularly relevant whenever the following occur: There is an initial statute or regulation implementation, changes in policies, or there are particularly complex scientific issues or controversial issues which need to be discussed.

Today's guidance documents involve areas identified for public discussion prior to implementation, and you have heard how this process will evolve. As previously mentioned, FDA is implementing new regulations concerning review time lines via FDAMA. We feel clearly articulated guidances concerning study design and data analyses are key to the successful implementation of FDAMA.

[Slide.]

Thus, our goals today and over the next two days are: to review both the general approach to the design and implementation of trials and disease-specific guidances; to present expectations in regard to protocol design, adherence to protocols, and monitoring of trial progress; to discuss the importance of the preservation of randomized groups; to

discuss subset or subgroup analyses, and to emphasize the importance of clinical goals in the plan for statistical analyses.

[Slide.]

This slide presents some of the elements of good study design. These elements include clearly articulated entry criteria, precisely defined endpoints, ongoing monitoring of study site performance resulting in real time adjustments, develop pre-defined analyses that have anticipated both the need for maintenance of the randomization and additional analyses necessary to define the effect of drug compliance and other relevant concerns.

[Slide.]

Another way of stating this is how to succeed in the arena of drug development once you have a well-designed trial, which clearly defines the population and endpoints to be studied, and the population needs to complete the study.

Well-conducted studies involving timely communication with investigators and sites to ensure the quality of the study's progress. When you are trying to be efficient, there isn't much room for sloppy work.

A well-monitored study, a well-documented submission, and well-organized submission, which is technically accessible.

[Slide.]

This morning, Drs. Chikami and Lin will further explore these trial issues. Additionally, they will review the need for reassessing how one sets the criteria for the difference determining ability of a trial.

These are trials with active comparator designs. We believe that the specific disease entity defines the difference that will be acceptable for a new therapy in comparison to an active comparator.

It would seem appropriate to us to place greater emphasis on the disease being studied and the consequences to the patient of a failure to successfully treat that disease. For example, we would not usually accept more than a 5 or 10 percent difference in cure rates.

You will also see this referred to as response rate or the delta, calculating the delta for the therapy of meningitis, but might accept a 15 to 20 percent difference in cure rates, or response rates or delta, for vulvovaginitis depending on other clinically relevant parameters, such as dosage form, prolonged half-life of the product, and the safety profile.

The therapy still may exceed the delta and still be a useful product.

[Slide.]

This approach is based upon the specific disease, and not just the response rate because efficacy is not simply a statistical goal.

Dr. Chikami will further discuss trial design and analysis issues, to be followed by Dr. Lin, who will place the statistical foundation for our journey.

DR. CHIKAMI: Thank you, Dr. Murphy.

As the third lineup of the introductory topic, I will keep my remarks relatively brief.

[Slide.]

I would like, first of all, to welcome everybody this morning, the committee members and consultants, and our audience who will participate in this process over the next three days.

I want to make a few general comments on the development of these guidances, and then I will touch briefly on a couple of issues that were raised by Dr. Murphy, but again, these will be discussed in more detail by Dr. Lin in her presentation of the statistical section of the general document, and also as the specific guidance documents are presented over the next couple of days.

[Slide.]

I think as was pointed out by Dr. Albrecht, the overall goal of this process of developing guidance

documents or the process within ODE IV incorporate the existing documents and newly developed guidance documents into a combined set of guidance documents that provide comments on the design and assessment of clinical trials for antimicrobial agents.

An additional goal is to provide clarification of previously issued guidances and to ensure that those documents are consistent with the current views of the agency on the general issues related to drug development and the reassessment of clinical trials.

[Slide.]

I think the scope of these documents include as reported general guidance for the design of clinical trials, antimicrobial agents, issues relating to trial implementation and monitoring, as was mentioned by Dr. Murphy, and finally, recommendations for analysis of these trials as will be discussed in more detail by Dr. Lin in her presentation.

[Slide.]

There are two general issues that I would like to introduce briefly. The first relates to the approach of the assessment of active control trials and how acceptable difference or delta is defined. The second relates to defining the patient population for the analysis that will

be conducted.

[Slide.]

When comparing an experimental therapy to a control, the comparison between the treatments is based on the 95 percent confidence interval around the difference. This was the approach that was first suggested in the Points to Consider document published in 1992. I think we will continue to apply this approach to the analysis of active controlled trials.

In the past, an acceptable result for the lower bound of this confidence interval was based in part on the success rate for the control arm. However, I think, as Dr. Murphy mentioned, in interpreting the results of such analysis, it seems to me it would be more appropriate to place emphasis on the disease being studied and the consequences of failure to treat it successfully.

[Slide.]

The assessment, however, should also take into consideration other characteristics of the drug, for example, the safety profile or potential dosing advantage. Thus, a drug for which the observed lower bound of the confidence interval exceeds the acceptable delta may still be considered a useful product, if, for example, it has fewer toxicities or, as I mentioned, the dosing advantage.

In such situations, labeling which describes these results may be appropriate.

[Slide.]

The second area I want to touch on briefly relates to the populations to be included in the analysis of trials. In analyzing these trials, the initial approach would be to keep the randomized groups intact as far as possible to protect the comparison of the treatment groups within a trial. This leads to what is often called an intent-to-treat approach, usually or may be defined as including all patients randomized in the trial.

I think there are situations, however, in which patients, based on predefined baseline characteristics, may not be included in such an analysis. For example, in anti-infective trials, the patient may not be included if there was a negative baseline culture. This is sometimes referred to as a modified intent-to-treat analysis and will be discussed further in the presentations.

[Slide.]

As Dr. Murphy mentioned, additional analyses may be performed to examine specific issues within a clinical trial, for example, issues related to compliance, missing data, discontinuations. These may be performed in addition to an intent-to-treat analysis to illuminate certain issues

that may arise in the interpretation of these data.

These should be prespecified including which subjects will be included in the analysis. We would expect that the results of such analyses will be consistent with the intent-to-treat analysis, however, it is important to remember that these analyses must be interpreted with appropriate caution.

[Slide.]

As I said, Dr. Lin will be discussing these issues in more detail in her presentation which will follow.

Just a general comment. As noted by Dr. Murphy, the process for the development of these guidance documents includes an opportunity for public comment prior to their implementation.

The documents have been posted on the internet and are available for public comment, and we would welcome and encourage submission of comments to the agency for our consideration and incorporation in modifying these drafts.

In addition, this meeting will provide important opportunity for us to obtain scientific input from our advisory committee members and consultants, and we look forward to the presentations and the discussions that will occur over the next three days.

DR. CRAIG: Thank you, Chikami.

Let me see if there are any questions about any of the comments that were made on the introduction from any of the members.

I guess the one thing that I would comment about is I hope we will get comment from the audience especially to bring up questions that you would like the committee to discuss, because you are going to have the opportunity obviously to submit your comments to the industry, but the committee is not going to be around to discuss those, so if there are specific questions that you think would be good for the committee to discuss, you will need to bring them up at this meeting, and I would strongly urge you to do that.

Without further ado, let's move on the to the first topic, which is Biostatistics, and Daphne Lin will give the FDA presentation.

Biostatistics

FDA Presentation

DR. LIN: Good morning. I am Daphne Lin, statistical team leader for the Division of Biomedics IV, Division of Anti-Infective Drug Products.

[Slide.]

Today, I am going to present statistical considerations for clinical trials in developing antimicrobial drugs. This is a subsection of the general

guidance for industry which we are proposing.

This is the joint work with Dr. Paul Flyer and Dr. Erica Brittain in collaboration with the medical divisions and other members of our statistical team.

[Slide.]

First, I will briefly discuss the topics included in the Statistical Considerations Section in the proposed guidance document.

Second, I will discuss the issues regarding the similarity, also called "equivalence" trials, especially choice of delta.

Third, I will discuss intent-to-treat and per protocol analysis, and discuss the issues regarding missing data, finally, some proposals will be made.

[Slide.]

This is an overview of the Statistical Section of the proposed guidance document. We have included sections regarding study design, data quality and management, and data analysis considerations.

It is impossible for me to discuss all of the topics which are included in each section. Some of the topics will not be covered today, for example, issues regarding sample size, entry data analysis, multiplicity of judgment, data quality, and data management will not be

covered today.

What I will talk about are the topics which I mentioned in the previous slide, like choice of data, intent-to-treat, and per protocol analysis, and the issues regarding missing data. These are the issues which may stimulate the most discussion.

[Slide.]

Next, I will discuss some regulatory background. The intention of similarity trials, also known as the "equivalence" trials is to demonstrate the drug is safe and effective. In 21 CFR Section 314.126, there is a discussion of what is meant by adequate and well-controlled studies.

This section describes five different types of control trials which may be used, one of which is the active control trial. The intent of the active control trial is to show similarity of the test and the control drugs, and active control trial is often used when a placebo is considered unethical.

It should be kept in mind that the lack of statistical significant difference will not be used as the evidence of similarity. Instead, a confidence interval approach should be used to evaluate the similarity of clinical effect.

[Slide.]

I would begin the discussion with a review of 1992 Division of Anti-Infective Drug Products Points to Consider document.

The Points to Consider document suggested to use a two-tailed 95 percent confidence interval around the difference in outcome to establish equivalence. For a cure rate of greater than 90 percent for two drugs, a confidence interval of zero, and lower bound data of minus 10 percent or less will usually be required to establish equivalence.

If the cure rate is 80 to 89 percent, the data will be minus 15 percent. If the cure rate is 70 to 79 percent, then, delta is minus 20 percent. In addition to excluding data as described, the Points to Consider also suggested that the confidence interval should also include a zero.

The Points to Consider also discuss some situations which may speak to the statistical definition of equivalence will nonetheless be clinically unacceptable, however, this special situation often overlooked by the sponsor, it appears that there is a tendency to realize whether the confidence interval include the delta regardless of the clinical situation.

This often lead to conclusion between the sponsor and the Medical Division which the risk associated with

treatment or failure is substantial.

[Slide.]

This shows the relationship between delta and the cure rate as discussed in the Points to Consider document. The choice of delta only depends on the cure rate. For a cure rate above 90 percent, the delta is minus 10 percent. If a cure rate of 80 to 89 percent, the delta is minus 15 percent. Finally, with a cure rate of 70 to 79 percent, the delta is minus 20 percent.

You should also notice that this is a step function which can lead to problems of interpretation. If a few values are changed, a different standard will be used for evaluation. For example, if the cure rate is changed from 80 percent to 79.5 percent, then, the delta will be changed from minus 2 percent to minus 15 percent.

[Slide.]

I have discussed that the delta has been chosen primary by cure rate, however, if it is important to choose delta to reflect important clinical factors, such as risk associated with treatment failure. The advantage and disadvantages of study drug, and historical cure rate with and without therapy.

In addition, if we have chosen appropriate clinically relevant data, then, there may be situations

where inclusion of zero is not necessary. For example, if the test drug has much less toxicity, then maybe we can sacrifice a little bit of efficacy. In such a situation, a large trial with an interval which is close to zero, but it does not include a zero, could still be considered persuasive.

[Slide.]

Therefore, we will have to propose that. When delta is chosen for sample size computation, it should be clinically relevant. Since delta will be picked based on clinical issues, it will need to be indication specific.

We are currently proposing indication specific to the recommendations. Of course, when making an indication-specific recommendation for delta, we will take into account the regulatory questions regarding delta which have been used to approve previous applications.

There are also special situations for individual indications where delta may be chosen on a case-by-case basis. For example, if a less effective control arm is used, a smaller delta may be required to demonstrate an experimental treatment is better than no treatment. We strongly encourage that.

Sponsors discuss the choice of delta with appropriate Medical Divisions during protocol development

stage.

[Slide.]

Since the similarity of the test drug and active control can mean that all drugs are effective or that neither is effective, therefore, we would like to recommend that sponsors should provide the rationale for selection of control arm to be used in the study.

This should be done at the protocol stage to ensure that the appropriate delta is chosen. This will avoid the concerns of the so-called "biocreep" phenomenon in which trials over time used progressively less effective control arms.

[Slide.]

Next, I am going to move to a different topic, the intent-to-treat principle. Many researchers define the intent-to-treat population as all randomized patients, however, some infectious disease trials have the complication that results are not present until after randomization. Our interpretation of the intent-to-treat principle is that it is permissible to exclude subjects based upon baseline characteristics. This approach is also known as modified intent-to-treat.

We have had the same confusions because the terminology of modified intent-to-treat has suggested that

we are not doing a real intent-to-treat analysis. Actually, the modified intent-to-treat analysis can be considered as a valid intent-to-treat analysis as long as the exclusions are based upon baseline characteristics and agreed upon in advance, then, we still have an intent-to-treat population.

Since there are a number of valid population, which could be called intent-to-treat, it is quite correct a precise definition be described and justified in the protocol. In the rest of my talk, I refer to intent-to-treat analysis while other presenters may call this modified intent-to-treat analysis.

[Slide.]

There are a number of advantages to the comparison of treatments used in the intent-to-treat principle. The first is that the comparison is protected by randomization. By this, I mean that groups are known to be comparable at the time of randomization.

Second, intent-to-treat can be interpreted as comparison of two strategies. For example, where patient assigned initially to arm A have an ultimate outcome different from those initially assigned to arm B, changes occurring after randomization will naturally be incorporated into the treatment comparison.

Instead of considering failure to take drug as

leading to non-evaluability, the intent-to-treat principle attempts to incorporate the information associated with compliance, however, a concern that is frequently raised about the intent-to-treat is that it may reduce the ability of the trial to detect a true difference between treatments. This is of particular concern in the similarity trial setting.

[Slide.

Because of this concern lends a desire to see if the exclusion of subjects with poor compliance and the missing data leads to the same conclusion as the intent-to-treat analysis, leads to per protocol analyses.

This type of analyses are sometimes referred as clinically evaluable analyses or microbiological evaluable analyses or clinical and microbiological evaluable analyses, however, the validity of the per protocol analyses rely totally upon the assumption that the two treatment groups after excluding non-evaluable patients are comparable.

In practice, there may be a selection bias. For example, treatment discontinuations may be related to the severity of disease. In fact, a per protocol analyses may be comparable to analyses event of observation or study, and it may be difficult to evaluate the bias due to a lack of statistical power.

Also, we may not always know the key variables or even have relevant information recorded. The absence of a statistically significant difference on baseline characteristics does not conclusively demonstrate a lack of bias.

[Slide.]

Both types of analyses, that is, intent to treat and the per protocol analyses are important for approval. The results of both approaches should be logically consistent. We would like to see that both intent to treat and the per protocol analyses demonstrate efficacy.

In the absence of such consistency, it would be the responsibility of the sponsor to provide a satisfactory explanation for any discrepancies.

Because the objective is to demonstrate a good treatment effect with both types of analyses, there is no need for a multiple comparison adjustment.

We also recommend to design protocol in the monitored trial to minimize exclusions, so that the intent-to-treat population and the per protocol population will be as similar as possible.

We also would like to emphasize that if there is substantial missing data and poor drug compliance, then, a trial's ability to demonstrate efficacy is weakened.

[Slide.]

Because of the practical difficulties associated with the conduct of a clinical trial leads always to a certain amount of missing data and noncompliance, therefore, each study report should contain a clear description of who is not included in the analyses and for what reasons with the comment of sponsors to submit the tables to account for status of all randomized patients for both arms, patients in intent-to-treat populations, and the reason for exclusion in the per protocol analyses, such as missing data and the lack of compliance.

[Slide.]

The issue of missing data is a problem for both intent-to-treat and the per protocol analyses. For example, for intent-to-treat, patients could have complied with the study medication, but the means, the test-of-cure visit, the protocol should specify preferred methods for dealing with missing primary endpoint. Many methods have been proposed to deal with missing data, however, each method has its own problems.

For example, in the intent-to-treat analysis, one method used to deal with missing data is to treat missing as failures, however, we know this could make arms look similar.

Another method is to assign outcome based on case-by-case review. This introduce a subjective component into assessment of treatment outcome.

Since each method of dealing with missing data has its own problems, it place us in an awkward situation. We would like to have a method prespecified for dealing with missing data, the adequacy of proposed approach which depends on the pattern of missing data.

[Slide.]

A sensitivity analysis should be conducted if there is missing data. We recommend that the sponsor should include a variety of strategies for handling missing data for both intent-to-treat and per protocol analyses.

The purpose of sensitivity analysis is to demonstrate that results of the protocol specified method do not inappropriately favor the experimental arm.

[Slide.]

In conclusion, we have the following proposals. First, delta should reflect risk associated with treatment failure. The advantage and the disadvantage of the study drug and the historical cure rate with and without therapy, that is, delta should be clinically relevant.

Second, design, conduct, and monitor trials to minimize missing data and the poor compliance to drug.

Third, both intent-to-treat and the per protocol analyses should be conducted with an eye to select both intent-to-treat and the per protocol analyses demonstrate efficacy. In the absence of such consistency, it will be the responsibility of the sponsor to provide a satisfactory explanation for a discrepancy.

In addition if there is missing data, sensitivity analysis should be conducted. In conclusion, the approval of a drug depends on many factors.

Thank you for your attention. I will stop here. We are very happy to answer any questions.

DR. CRAIG: Any questions specifically on the presentation, either topic covered?

I guess the only question I would have is what happens if the efficacy is less than 70 percent, what kind of delta do you use then?

DR. LIN: In the past, they use minus 20, so, for example, like for a drug product, you know, the cure rate, sometimes cure rate will be as low as 60 percent or 55 percent, and which they use minus 20.

DR. CRAIG: Thank you.

Use the mike and introduce yourself. State your name and your organization.

DR. JOHNSON: My name is Roger Johnson. I work

for Zeneca. I am a statistician there.

I just had a question, on the confidence interval you said you use a 95 percent confidence interval for equivalence trials, and I have been using 90 percent confidence intervals although not for the case of binomial data, but it's equivalence trials establishing for continuous data, so I was just wondering why you use 95 percent confidence interval rather than 90, because I thought the 90 was standard.

DR. LIN: No. You know, other Medical Divisions, they also use 95 percent confidence interval. This is--I don't think it is appropriate to say policy, but--

DR. ALBRECHT: In new drug evaluation, I believe 95 percent confidence intervals are used. Ninety percent confidence intervals are used by the generic drug groups to evaluate bioequivalence trials, and we have, in fact, in the vaginal drug area, in the study of generic drugs where, of course, bioequivalence is on a clinical basis, Dr. Winfield has had a lot of experience reviewing those studies, and the statistical analyses of those studies did employ the 90 percent confidence intervals. Maybe that is what you were alluding to.

DR. JOHNSON: I know that 90 percent confidence intervals, it is equivalent to the two, one-sided test

procedure where the alphas on the one-sided tests are each at 5 percent.

DR. LIN: We are using 95 percent confidence interval, so one side is, you know. 0.025 instead of 0.05. Like I say, you know, bioequivalence, we use the 90 percent, two, one-sided tests.

DR. CRAIG: We have Dr. Wittes, who is going to give a committee presentation.

Committee Presentation

DR. WITTES: Well, I don't know if this should really be called a committee presentation, it is officially, but it is really my views of what has been presented.

First, I really want to thank the FDA for inviting me here. I think that what you have all done has been terrific. I was very enthusiastic when Daphne and Erica and Paul called me. I was less enthusiastic when I got this book, but, nonetheless, I think that by going through in a very systematic way, thinking about the issues in their generality, and then trying to apply them to each disease entity has been very useful, and I think promises really to make things very clear and rigorous in the future. I congratulate you all.

I also want to say I love the image of biocreep, which I had never heard before, but the idea of this amoebic

blob moving backwards and backwards and getting worse and worse is just marvelous.

I think the main issues that have been brought up in the Statistical Section of the general document are the issue of missing values, how to handle that, the issue of equivalence/active controls, and I think the question that came on the floor, the difference between the 90 percent and the 95 percent confidence interval addresses one of the problems about how we think about trials that are comparing active controls.

The issue of intent-to-treat and the language attached to that, as well as which of the groups that should be looked at, and the difference between an intent-to-treat kind of analysis and the per protocol analysis, and then there were some other issues that I would like to bring up if it is okay, that were in the document that Daphne specifically said she wasn't going to talk about today, but just to address some issues of randomization, multiplicity, and interim analysis, they were touched briefly in the document, and I would just like to say a few words.

I would first like to talk a little about the language, using the intent-to-treat language, I think one of the confusions that comes up with intent-to-treat, and modified intent-to-treat, and maybe it's intent-to-treat,

and almost intent-to-treat, is that what we have got is really one population that is a population that the study group to which the randomization has occurred, and then there is another group, and the language of the document is very nice, because it speaks of the primary analysis population, and I think that is a very useful concept, that rather than say my intent to treat is better than your intent to treat, we have a group of people that were randomized, we have a rational, rigorous definition of what the primary analysis population is.

That analysis population is based on baseline exclusions, not post-randomization exclusion, and therefore, it is a rigorous analysis, and I think that would avoid some of this intent-to-treat language that can be very confusing.

Then, we can face the question about whether an exclusion is, in fact, a baseline exclusion, so that an exclusion, one of the reasons that excluding on the basis of the organism that plates out, which really occurs after randomization, is because the organism was there before randomization.

So, you can think in terms of what was the state of the patient at the time of randomization rather than when the measurement was taken.

Now, obviously, in the situation where there are

no missing values, and that was all the nice motherhood words about how it is good to have no missing values, then, we don't have many problems about the distinction between the analysis as planned that reflects the randomization and the per protocol analysis, but unfortunately, in the real world, there are very often missing values.

Some of the missing values stem from sloppiness of the study, and whatever any kinds of mechanisms that one can institute to prevent that, one should do.

There are, however, missing values that come out, that are inevitable in complicated studies, and so the question then becomes how best to deal with them. The thrust of the document and some of what Daphne discussed was that there should be two analyses, what I am now calling this primary analysis and then a per protocol analysis, and that the two should be logically consistent with each other.

Now, again, I think that is going to be fine as long as there aren't very many missing data. I mean then it will be a surprising situation in which they were not logically consistent.

In the presence of a lot of missing values, I believe that what we really have to acknowledge is that I think I would actually go a little farther than what Daphne did, that, yes, it is important to do, in fact, necessary to

do sensitivity analysis, but I think that there comes a point in the study where one has to say what one has is an epidemiologic comparison. One started off doing a randomized trial, one has lost so much data that one really can't fix it and still call it randomized.

I think that the problem what we talk about is statistical concerns about missing data, they are really medical concerns, as well. I mean do we have two comparable groups and can we compare them in an unbiased way.

There are at this point in the commercial, I mean any statistician opening the mail now gets all kinds of advertisements for how to handle missing data, computing techniques to handle missing data.

I think we need to be very careful, again acknowledging that these are essentially epidemiologic analyses that adjust for failure of randomization because that is what the effect of large amounts of missing data are.

A couple of words on equivalence. It is nice to see the moving away from the step function analysis. I think that led to lots of logical inconsistencies and led to, I think, an encouragement for sloppy studies, because you gained a lot, a little bit of sloppiness gained to a huge amount.

I think it is also very good to not demand that zero be in the confidence interval because once more, demanding that zero be there, as you are looking for wide confidence intervals, and the tight confidence interval that excludes zero at either end is more informative than a wide confidence interval that includes zero.

I think again part of our problem is just like intent-to-treat has been a problem word, I think equivalent is a problem word, because when we talk about something being equivalent, it is hard to say, well, it is equivalent, but we know it is not the same.

So, I think again the use of these words active comparative trial, getting away from language that forces us to think about these things must be the same, I think is very useful.

I would like to hear Bill's comments about the use of the confidence interval approach because there are other approaches that people have used in dealing with these kinds of trials, and although I find the confidence interval approach very easy and interpretable, I think it is important to open it up and talk about other ways.

Back to the business about sensitivity analysis. I think prespecification of the approach to sensitivity is necessary, but not sufficient, because one can prespecify a

narrow enough range of sensitivity analysis to get the answer that you want.

I think the idea of sensitivity analysis is that you ought to be analyzing in as broad a way as is medically sensible, that one doesn't need to necessarily go to the most extreme, worst-case analyses, but the worst, reasonable-case, over a range of kinds of assumptions is what one wants to do, and that is hard to prespecify because sometimes the pattern, as Daphne pointed out, sometimes the pattern of missingness is much more complicated than one would have predicted.

You might have a small amount of missing data with a disturbing pattern, and that can lead you to really try to explore the data and make sure that what you are seeing--what you think you see--is likely to be there, even have the data not be missing.

A couple of issues that Daphne did not discuss, but was brought up in the document itself. First of all, one of the things that the document warned against was complicated randomization schemes, and I just want to put in a plug for maybe we should consider them.

I have been personally very loathe to use anything but very, very simple schemes. In the very recent past, like last week, I was convinced by somebody for a very

complicated trial that it was really important to stratify by three or four different variables, and I resist stratification usually at baseline randomization, but I got convinced, and I felt that the only way to do it in this particular trial was to do exactly the kinds of randomization that you are warning against, where you are doing dynamic allocation.

I open it up as a question to you, to tell us why we shouldn't do this. I mean this is actually the first time I am venturing into it, but it seems to me that there should be times when if you really feel that there are baseline variables that are very highly predictives, there ought to be ways of balancing.

I would like to see more discussion of multiplicity. I think this is an issue that ties many of us up, it is very complicated, and the approach that the document took was to use a simple method like Bonferroni, and the justification was that you then can get uniform confidence intervals. I know that there is a movement toward using much less conservative approaches, and I would like to hear and I would like to see some discussion about it in the document.

Finally, actually, not finally, next to finally, the penultimate thing is the interim analysis. There was a

very short description of interim analysis that focused--it basically said you should only do it when you are talking about mortality.

My own feeling is that it would be nice to have a little bit more latitude, that there are sometimes other kinds of endpoints I think in which one would want the ability to stop early. Again, it is always an issue of whether there is an update for safety, and that may be where you are coming from.

Finally, the biocreep, and this really is the finally. I realize from reading the document exactly how important this issue is, and for me it was a very new idea, but the question comes up what is the sponsor to do.

What the guidance says is think about it, but it doesn't really help that much. It seems to me that there is two logical approaches that I could think of, neither of which were too satisfying.

One is use only a comparator that has itself been compared to placebo, and that I could see could be very difficult because you are taking something that hasn't been used in years, or the other approach might be to only take as a comparator one that has at least shown a directional favor, that it's directionally better than what has been shown to placebo, so sort of ratcheting away from the

biocreep, but again I would like to see more discussion of the way in which that decision should be made.

In conclusion, I think this is a great--oh, yes, one more thing, sorry. What I did notice in the documents as we went through this huge book was that there were, although the statistical guidance talked about the primary analysis as being the one that reflects the randomization, and the secondary analysis being the per protocol analysis, there were reversals in at least one of the documents, and I really think that when there are such reversals, there has got to be very, very clear justification of why the agency feels that the other analysis is primary.

Committee Discussion

DR. CRAIG: Questions for Dr. Wittes?

I guess I would have one. I heard you mention a lot that here you were talking about missing data as a small amount of missing data, too much missing data. Is there any way that one can quantify what is sufficient to miss and what is too much?

DR. WITTES: Would that there were. I guess the rule that I kind of use is if there is so much missing data that biologically and medically reasonable conservative assumptions would change either the direction results or the magnitude of the effect in an important way, that is too

much, so it is kind of a "you know it when you see it, but it is hard to specify up-front."

DR. CRAIG: Difficult for non-statisticians to understand.

DR. GESSER: Richard Gesser, Merck Research Labs.

Regarding baseline cultures, our cultures on enrollment study, it wasn't clear to me whether you considered those baseline values or not. Those are cultures whose results are not available until after enrollment.

DR. WITTES: Oh, I do, absolutely, because they characterize the patient at baseline.

DR. GESSER: Fine, but even though that information wasn't available to the investigator until subsequent to the study, fine.

DR. WITTES: Yes.

DR. GESSER: Thank you.

DR. CRAIG: Dr. Murphy.

DR. MURPHY: I think one way of summarizing is when you are using the culture, you don't want to study people who don't have the disease. That basically is why we think of it as a baseline characteristic, that they have to have the disease that your culture is. It is just a matter of infectious disease, we have to deal with this as a criteria that doesn't come up or present itself until days

later sometimes.

DR. CRAIG: Did the FDA have any comments? Yes.

DR. GOLDBERGER: You were talking about stratifying during the randomization. Were you also referring, just doing that for purposes of balance or were you speaking in terms of actually doing analyses by those particular strata?

DR. WITTES: I was really just speaking of balance. I brought it up because there are methods to do stratification to achieve balance when you are not stratifying--let's say you have three different organisms and 10 sites, 10 clinical sites. So, you block your randomization with any site, and that would be 10 strata.

Then, there is these three organisms, and you might want to randomize within, so that would be 30 strata, which is a lot of strata. So, there are methods where you can say what you want is balance across the entire--you want a third of each of the organisms in each of the treatment groups across the study, and those are done technically in a dynamic way.

I mean you have a central randomization, and it balances across the study. You don't have to send in kits and blocks. Those are complicated to do, and my reading of the document was that you guys were discouraging that, and I

just--maybe I am wrong--

DR. GOLDBERGER: Let me just ask you one other question. I think one issue also comes up whether a clinical trial for most of the indications would be sufficiently large that the cells that you would produce by that type of randomization would have a large enough n that they would be useful. I am just not clear on that.

When you were referring to using that technique, I don't know the size of the trial that someone had convinced you. Some of our trials, as you know, are not really that large.

DR. WITTES: The trial that I was dealing with is a trial of 400, which is not all that large, about 10 sites, and two variables that the clinicians really want balance in the treatment groups. So, I don't believe it is possible to do the analysis, to do so many strata in the conventional way, I just don't. I think the cells will be too small.

Now, I think the issue of whether you analyze it by the strata or not, I think that is a different issue.

DR. GOLDBERGER: But that wasn't your intention.

DR. WITTES: That was not my intention.

DR. GOLDBERGER: My other comment was just very brief. When you talked about how you quantify how much missing data is too much, is another way to simply say it

that if the amount of missing data is small compared to the confirmed endpoints, so that you are unlikely, as you said, to disrupt the conclusions of the analysis, I am trying to think of a way that it is easy for people to grasp in terms of beyond just saying you know it when you look at it.

DR. WITTES: There is one thing that people do is to look at the cohort that is missing and ask how they compare at baseline to those who were present, and they say that if they look the same, well, you don't have to worry about it. I don't like that at all.

It seems to me in that case, even if you have a small amount of data, but if the cohort looks different, then, I would worry. If the cohort looks the same, then, I would start asking questions about suppose doing these aggressive sensitivity analyses, making strong but not crazy assumptions about the direction, about how the endpoints would have fallen, might have fallen had they been observed.

Only if those don't change the qualitative results, and the qualitative results including the strength of the estimate, then, I would say yes, that's fine, the missing doesn't matter very much, but if it could change anything that makes you change your way of looking at the results, then, I would worry.

DR. FLYER: Hi. I am Paul Flyer. I will try to

address some of the statistical issues I have been working with. Daphne basically flipped a coin to see who would present, so whoever presented didn't have to deal with the questions.

I was hoping I was going to make it through without having to deal with these issues, and they would just sort of sail through, but I guess that is not to be.

The randomization is a very complicated issue, and I would be very reluctant to separate analysis from the stratification because we have a long-standing principle of telling sponsors you have to analyze as you randomized.

Now, once you introduce the complicated method of stratification, you have to reflect it in the analysis and usually the gains of the stratification aren't worth the cost of the analysis.

I am very, very loathe to bring up very complicated statistical issues in an advisory committee and say, well, gee, if you analyze it this way you get this result, if you analyze it this way you get this result, and you don't want to do a complicated analysis because you have complicated designs, so you do a simple analysis, but a more complicated one leads to a different result.

It is very awkward and it is counterproductive, so that when you start to get 30, 40, 50 cells, your gains on

the stratification will be minimal once you get past a couple of cells, what you are really achieving is a cosmetic balance.

So, what we are trying to tell sponsors is if you use a complicated design, it ought to be justified based upon your variance reduction, and when you do your analysis, you have to think it through and figure out, well, how do you reflect your variance reduction in your analysis.

Even with a simple, highly stratified design, not using a dynamic allocation, which I will get to in a minute, you still end up with lots of complications. Sponsors say, well, we didn't have enough people in each cell, how do you do a stratified analysis, and we get into lots of arguments, and that stratification was unnecessary because if you have a lot of small cells, you are not getting big gains in variance reduction.

So, what we are trying to do is encourage people think through the analysis, how are you going to want to show this, and since you are going to be obligated to reflect what you have done in the design, think it all the way through including the analysis.

Now, when you get to the dynamic allocation, no one saw through the analysis, there is not a literature which says this is the correct analysis based upon, for

example, the randomization perspective or even a good model-based perspective, because it will depend on assumptions.

I don't want to come to a committee and discuss statistical assumptions and whether they are important or not, it is more important I think to focus on the clinical issues, the missing data, what have we actually seen rather than do esoteric statistical arguments on how we did our randomization-based analysis.

So, I think unless there is a great gain in precision by the proposed design, stick with the simpler design.

So, that is the principle which means anyone who wants to propose a complicated design, thinks it is justified, should go ahead and use it, but then know they are obligated for the analysis and how to reflect that in a way that people will agree on.

If it gets sort of messy after the fact, let's say you have a result that is very close to the magic 0.05, it is awkward if someone else says 0.07, 0.8, then, 0.03, and since there is not a clear agreement in the statistical community on how you analyze a complex design, that is awkward, and the sponsor should think that through beforehand. That is why the advice is keep it simple unless

you have a good reason for doing it otherwise.

Similarly, with Bonferroni, the confidence interval adjustment, a lot of the more complicated procedures require just that, complicated decision rules. Well, what we want is to be able to clearly understand what are the confidence intervals showing.

Some of the hierarchical procedures don't really even have simultaneous confidence interval interpretations. So, that gets a little awkward because you end up having to specify basically a testing procedure, and that can get awkward. It can complicate some of the analysis.

For example, some of these methods say that unless an overall result shows something, you can't look at the individual results. Does that mean that if we have two trials, both three arms, one trial says we can't look at the individual comparisons because this sort of global test didn't work, but then another procedure for another trial shows the result, do we have two trials still? But there could have been enough evidence in the first, so that it could have been useful.

So, I think based upon sort of the glazed looks in the room, it is not the type of discussion you want to get into at an advisory committee, so it is sort of unless you have a good reason, keep it simple.

I think interim analysis is another good statistical thing that we are interested in, but there is another concern, which is we want to have sort of a database, for example, 500 followed to six months, or a 1,000 followed to six months.

The interim data analysis for a strong result could stop a trial early, before we really have enough information to characterize, let's say, the safety profile of the drugs. We have said in public that unless there is irreversible morbidity or mortality associated with the endpoint, there may be more of an incentive to let the trials go to completion, so that in situations where the outcome is so severe that we really want to know early, then, you might be willing to sacrifice a little safety information because the efficacy result is so important.

So, that was where the recommendation came from in the document that interim analysis be avoided except in extreme situations in the interest of having trials go to normal completion.

Biocreep, I think I will leave for the Medical Division, if they want to deal with that in terms of picking an active control. The ITT is primary. I think we have been trying to avoid the designation of something that is primary because it leads to lots of battles over sort of

prespecification primary - no, I want this to be primary, no I want this to be primary, where, in fact, you want to look at the data as a whole and judge the results, and avoid unnecessary arguments about, well, what is technically primary when we know we are going to want to see consistent evidence over the analyses as a whole.

So, we haven't found these are primarily always productive topics for discussion.

Do you have further comments on that?

DR. WITTES: Yes. Actually, I agree with almost everything you say, and I think it would be good to expand some of the discussion in the document because if your justifications had been there, then, I think it is very clear where you are coming from. This way, there was a kind of oh, my goodness, I can't do that. Why not?

DR. FLYER: There was an edit. We started out with more, and went through the editorial process, it was pared back because it was thought that the document was just getting sort of unwieldy, so that we started with a lot of this detail, and it was pulled out the last couple of versions. Maybe we can beef it up a bit more.

DR. WITTES: Creep it back in.

DR. CRAIG: Dr. Blackwelder.

DR. BLACKWELDER: Thank you. I would like to

thank the FDA for inviting me also because I think there are some important issues in an area that I am particularly interested in, so I would like to make a few comments. Mostly they are emphasizing points that have already been made by Dr. Lin or Dr. Wittes.

I will start with the one that to me I think is the most important, and that is from the 1992 Points to Consider, the idea of the confidence interval including zero. I think that is not a good criterion because it discourages too large a study, and I think that is the point that Janet already made. It kind of says don't look too hard for a difference.

Further, the fact that a confidence interval is zero does not give us confidence that there is not any difference. I think that was kind of the implication I got from the criterion in the first place. It doesn't tell us there is no difference or that two drugs are equivalent, and that leads to the language problem again.

Equivalent is probably not a good word to use. I like similar, but there is something better than equivalent because we are never showing the two preparations are equivalent, I think.

Ideally, in this kind of design, you would select, would have this delta, this difference we want to rule out

that is both clinically meaningful in the sense that as long as the difference is less than that, we are willing to use the new drug, not calling it equivalent necessarily, but saying we are willing to use it, it's similar enough.

Sometimes that can lead to large sample sizes. I know depending on the context, and then I am not sure what the best idea is. One possibility I have thought of that would need further study, I don't think it is ready for a recommendation in a guidance document, is to require something like a point estimate of the difference be less than delta over 2, if that would be meaningful.

As I say, I think that needs some more study, but it is something I have played with a little, and I think might be explored.

One other point about the delta that I think was implied in the presentation that was made is that besides being clinically meaningful, it has to be small enough so that once you are through, and you show a difference less than delta, you are sure that the new preparation is effective. In fact, it should be clear from the trial that both are effective in this particular trial.

A couple of other comments about the 1992 Points to Consider. Again, the 95 percent confidence interval, I think we need to recognize--and I just want to underline

what Dr. Lin said in response to the question--that when we use the 95 percent confidence interval, we must be clear that the error rate that we are using for the question of the difference being less than delta is really 2 1/2 percent, so another approach one might take to this, although I like the confidence interval approach, but you can think of it as testing a hypothesis, and if we are testing a hypothesis with a 2 1/2 percent alpha rate, I am not saying that is good or bad, but we just need to recognize that is what it is.

I also like the idea of not having specified deltas for all studies, just depending on the risk or so-called cure rate that we are talking about. I think there are a lot of factors, most important of which is seriousness of the outcome probably that need to be considered when coming up with an appropriate delta/delta that is clinically appropriate.

I would like to make a comment about the intent-to-treat analysis in the context of a trial to show similarity. It has been pointed out in some papers, and I think it is pretty clear that in that kind of trial, the analysis of all the randomized individuals might, in fact, give you the result you are looking for when it is not true. In fact, it can inflate the so-called Type 1 error rate.

So, it may not be clear what the most meaningful analysis is in that context, if there is a lost of missing data or a lot of noncompliance, which is a way you can get things to look similar when they are not, then, the study just might be in trouble, so it is incumbent on the investigators, I think, to be very, very careful, maybe even more so with this kind of study, to see as much as possible that patients are compliant and that data are not missing.

For the biocreep--I like that word, too--it may not be possible or feasible, as Janet has said, but as much as possible, the only solution I can think of at the moment is to go back to the original active control or as close as you can get and to have a trial that is designed as similarly as the original placebo-controlled trials as you can get, or if you are fortunate enough to have a treatment that has been shown to be better than the active control, certainly use that, as I think Janet suggested.

Those are all the comments I have for the moment. Thank you.

DR. MORRIS: David Morris with Abbott Laboratories. I would like some clarification of another problem, which is consistency, which I have heard mentioned several times in terms of consistency of intent-to-treat and per protocol.

I am curious, will that be interpreted as the confidence interval criterion is met with both data sets or is it a more directional sort of approach.

DR. FLYER: It is one of those it depends a lot. I don't think we were interested is saying that both intervals have to make it because of the problems with the missing data could suggest after we do the review that we have more confidence in one than in the other.

I think what we are looking for is a good basic consistency both supportive if they are multiple trials being used to support the indication, that those other trials also be supportive, I think, with superiority trials the same principle would apply that one analysis makes it 0.06, the other one makes it 0.04, is that a failure or is that basically confirmation that the result is reasonably robust.

I would interpret the latter as that those results are supportive, so I think that is what we will be looking for, that if the one that fails is failing in a way that we don't have confidence in because of the pattern of missing data, that would be the one we would tend to dismiss, so we are trying to be flexible and base it upon a review of the data and the overall pattern of evidence.

DR. CRAIG: Yes, Dr. Wittes.

DR. WITTES: Let me ask you a question. In a situation where you are actually anticipating a lot of missing values for perfectly legitimate reasons, there is a structure where people are going to not comply, and you know that, where would you size it? Would you size the study for the less powerful of the two?

DR. FLYER: That would be prudent, but power and sample size is really the sponsor's risk. I think the recommendation would be from us that it would be prudent to power up for the analysis that you think you will be weakest on, but the sponsor has to decide to what level of risk are they willing to assume, but it is also what other trials would be submitted, so that I think a relatively weak result in one trial could be made up by a very strong result in a cleaner, easier to run trial.

It is a hard question to answer. I don't know if Gary or Dianne or Mark want to chime in. It is difficult knowing these are statistical issues versus the Clinical Medical Division, it is a fuzzy sort of area.

DR. MURPHY: Obviously, if you had a highly effective therapy, you might be willing to take the risk as far as that is what you are balancing, right, is the size and the cost and the ability to complete that study well versus which group would be your analysis, final analysis

that you would be looking at.

I think that, as Paul said, that really is something we can't give a formula for, that it is going to have to be looked at with the activity of your drug, the available population, and the ability of your investigators and sites to deliver what they are supposed to.

DR. FLYER: With the HIV drugs, we have been trying to change trial design as much as possible to match the medical practice as we understand it, so that if people can't take a drug, and that is an important consideration in the patient's well-being, that that could be treated as a treatment failure, not just as a missing data technique, but actually treatment failure.

So, in situations where compliance is a real problem, and it is part of the medical situation to try to sort of fold that into the interpretation of the trial results as part of the endpoint whenever possible, but then we have to make sure that that is not being manipulated in such a way to make a lesser drug look more effective. It becomes a difficult issue, but try to make the trial design reflect clinical practice as best we can even though we know trials can never really mimic in clinical practice, but I think a sensible trial design can hopefully avoid having to try to fix things after the fact, which none of us will ever

be satisfied with.

DR. CHU: Ray Chu. I am a statistician from Rhone-Poulenc Rorer. The first question is regarding just the name of equivalence claim. Here, it might be a little misleading based on if we only use lower confidence limit, that sounds more like non-inferiority for the objective of the test.

Another question is should we consider other type of measurement of deviation when we assess equivalence. Here, we are trying to use the difference between success rate, and some presenters also mentioned that delta should be not response rate dependent, but actually when you look at a different range of the success response, the failure deviation really has different impact imposed on the population.

For example, 95 percent success rate, if we allow, say, 85 to be equivalence, the failure rate actually increased from 5 percent to 15 percent, which is three times, whereas, if we look at 75 percent range, even drop with delta 20 percent to 55, the failure rate increase is less than two times.

So, I think maybe we should still consider the different range of success rate when we consider what is the meaningful delta. Of course, I agree we should look at

delta for a given type of infection, a serious infection, we really should consider the practical implication of really feasibility of a conduct study.

That is my question.

DR. CRAIG: Any comments?

DR. FLYER: I guess I heard more that there was basic agreement to what we are doing, that, in fact, when we consider going from a 95 percent success rate to 90 percent, we would take into consideration going from 5 percent failure to 10 percent failure, a doubling, that as we work through what we are trying to accomplish with the specific indication, I think your comments will be taken into account.

DR. WITTES: I think one of the important things is to stay away from the step function. The idea of it was to do exactly what you were saying, but not in the step function way.

DR. CRAIG: Would that difference be put into the label at all?

DR. FLYER: How do you mean?

DR. CRAIG: For example, if you had 95 percent for the control and 90 percent for the new agent, would that still somehow get into the label, or are we still saying they are similar?

DR. FLYER: How are we going to handle labeling?

DR. GOLDBERGER: I think that is something that often goes on a case-by-case basis. Sort of judgment is made as to how different the products performed, taking into account again the severity of the illness and the consequences of failure, and it may be appropriate, for instance, to put something like that in the Clinical Studies Section, even if it has, for instance, met what criteria were agreed upon before the clinical trial began, and certainly if it is a situation where one might not have met the agreed-upon criteria, but after the fact, taking into account some of the other benefits of the experimental arm, it seems prudent to prove it, there would still be probably some qualification perhaps in the Clinical Studies Section.

DR. CHU: I would just quickly add has like the measurement of odds ratio been considered instead of just using the response rate.

DR. FLYER: Right. I think it is the same idea, will reflect the difference in odds ratio and the delta is something that we would be looking at, as well.

DR. BROWN: Mike Brown. Bristol Myers Squibb. I have a lot of questions actually about the randomization and analysis, this randomized issue. I agree with Paul we probably need a separate forum to discuss this, because, as

you know, it is not just an issue in this division, but probably across all divisions, but more generally right now, perhaps you could address often when we do stratification or balancing within a trial, it is not so much for variance reduction because we think the factors are prognostic, but it is more for logistical considerations particularly, for example, within sites because of drug distribution issues and things along those lines, hence, we want to keep a balance just so it is easier to ship things out.

I was wondering your feelings on that side of things.

DR. FLYER: I guess what I tell sponsors is that if the analysis that you are using will be approximately reflective of the actual randomization, that would be appropriate.

Usually, that is not in the sponsor's interest because there is some variance reduction associated with the stratification, so if the sponsor proposes based upon what they know about the design, the factors that are likely to be a conservative analysis, and the sponsor feels it is in their interest to submit a conservative analysis, I wouldn't object to that as long as there can be some demonstration that wouldn't be inflation of Type 1 error. I wouldn't object if they wanted to do that.

DR. CRAIG: Any other questions? Does the FDA have any questions from the committee or any additional input, or do you feel you have gotten some useful comments?

DR. MURPHY: We definitely have gotten some useful comments, and I want to thank our consultant, that many of the important issues that we wanted brought out today have been brought out, and I want to thank the speakers also from industry, because these are issues that we are dealing with and we are shifting how we look at wanting our trials completed, because the point that was brought up over and over again is that missing data is a big issue, and we cannot be sitting here and deciding post hoc who goes where, what, or how you define them, are they a failure or not, and these are issues that we want to have this type of discussion and will be taken into consideration along with the comments, which, as we indicated, we are open for comments for the next 90 days, and we look forward to seeing those, and we will be publishing further information on this after we receive those comments.

Thank you all very much.

DR. GOLDBERGER: I am just curious about one thing. As Dianne said, we are making some changes in how we approach things and particularly emphasizing the seriousness of the disease and the consequences of failure in looking at

the delta as opposed to simply looking based on response rates at the control arm.

There is a large number of people from industry in the audience. We have certainly heard some comments from statisticians from industry. We haven't heard any comments from any of the clinical side in industry.

I was wondering if anyone had any comments, discomfort, et cetera, about this approach.

DR. CRAIG: If no one is going to be brave enough to stand up, please address that in your comments from industry back to the FDA, so that those can be addressed.

DR. BLACKWELDER: Could I make just one other comment? I might have misunderstood what Paul was saying just a moment ago about stratifying randomization by site. In my experience--and this is when we have sites that have a fairly large number of patients from each site--we do it as a matter of course, always stratify, mainly I think because we can't assume that populations at different sites are the same. So, I am not sure, I was surprised if the recommendation is to discourage that, and I might have heard wrong.

DR. FLYER: No, I wasn't trying to discourage that. I guess I was being somewhat cryptic, that if the sponsor wants to submit a conservative analysis, I was

trying to convince them through that argument that is not in their interest, that as a matter of course, you should reflect the design because you have chosen the design to generally minimize variance, and even if it is done for administrative convenience, it is usually things related to the outcome, like centers usually relate it to the underlying rate of the disease, so that you will generally be ahead of the game if you reflect the design through stratification of the analysis, but it gets a little frustrating going over and over the same arguments about whether you have to do center adjustments or not. We try to convince them it is in their interests to do it.

I guess I am just wearying of the debate a bit because usually the p values are just slightly smaller when you do the appropriate stratified analysis, but there are other issues I would rather deal with, like the missing data problem, which I think are much more important than that sort of issue, I think good design and good analysis would be important in stratification of the design and the reflected in the analysis, I consider to be appropriate and a good thing to do. It just gets a little tiring having the same debate over and over again.

DR. CRAIG: If there is no more, we will end this. You have got your first homework assignment, those from the

industry, is at least to have your clinical people comment on the statistical approach that is being used by the FDA and include that in any of your comments that you submit back.

We will take a break now until 10:30, at which time we will start on bacterial meningitis.

[Recess.]

DR. CRAIG: The first clinical entity that is going to be discussed was not presented at the previous session. This is on bacterial meningitis.

The FDA presentation will be given by Dr. Rakowsky.

Bacterial Meningitis

FDA Presentation

DR. RAKOWSKY: My name is Alexander Rakowsky. I am a medical officer in the Division of Anti-Infective Drug Products. The presentation will involve acute bacterial meningitis.

[Slide.]

This morning's entertainment will be provided by me as far as general overview of guidelines. Dr. Chesney will then present a summary of several key issues that were discussed in these guidelines. Then, hopefully, we will have a lively discussion of questions or any concerns from

the floor.

[Slide.]

As stated, this indication essentially deals with acute bacterial meningitis, meningitis referring to infections of the linings of the CNS. The next two slides will deal with the bacterial and acute sections.

The bacterial, in regard to the pathogens most commonly seen, the big three historically have been Haemophilus influenzae, Type B, which has dropped off considerably in this country, but it is still a problem overseas, and then presently Strep pneumo and Neisseria meningitidis in this country as the two most common causes.

Depending on the age group, you can also see other pathogens, Group B Strep, escherichia coli, and listeria monocytogenes can occur, with listeria also presenting itself in the elderly.

[Slide.]

Again dealing with the acute manifestations of this illness, there are four infections due to the following causes will not be addressed by these guidelines. I want to start this slide off by basically stating that the study of these is encouraged, but will not be discussed at this time.

So, infections secondary to in-dwelling catheters involving the CNS, infections in patients status/post recent

neurosurgical procedures or craniofacial fractures or trauma, anatomic defects predisposing to CNS infections, and then immunocompromised patients where mycobacterial, fungal, parasitic, or viral infections are either seen or strongly suspected.

[Slide.]

I just want to discuss several recent developments first before getting any further in these guidelines. The first and most important is the development of the HIB vaccine and the widespread use of it in this country, which has led to a rather dramatic decrease of, one, acute bacterial meningitis overall, and two, HIB-associated meningitis.

This is a good phenomenon as far as clinicians and patients are concerned, however, both the epidemiological shift more now towards Strep pneumonia and Neisseria, especially Strep pneumonia having higher mortality and morbidity rates historically, there is a possibility that for those patients who will now be enrolled in studies, we may actually see higher mortality and morbidity rates than seen in virus studies.

The second recent development is the use of dexamethasone. Dexamethasone was almost universally accepted for use where Haemophilus was suspected as a

pathogen, and there are several elegant studies looking at its potential decrease of hearing deficit status post meningeal infection.

The data is probably less convincing for Strep pneumonia, and there is very little convincing data for Neisseria, which brings up two issues, one, for studies done in areas where Haemophilus is still an issue, dexamethasone is routinely used, and may not be routinely used in this country, as Dr. Chesney will discuss; and, secondly, dexamethasone did decrease morbidity rates especially hearing when involved with H. flu. Now, with H. flu being enrolled in studies because of the HIB vaccine, there is potential less impact of dexamethasone and decreasing morbidity, yet again leading to potentially higher morbidity rates in future studies.

Lastly, Strep pneumonia resistance, there is an exponential increase of non-susceptible to both penicillin and cephalosporin strains of Strep pneumoniae. In some parts of the country, it is approaching 40 percent. Again, Dr. Chesney is an expert in that because of the problems in Tennessee.

There was theoretical risk of having treatment failure secondary to decreased susceptibility to third-generation cephalosporines, and these were basically

verified by multiple anecdotal reports.

That led to the American Academy of Pediatrics in the February 1997 issue of Pediatrics to present guidelines about the empiric use of vancomycin in situations where acute bacterial meningitis is suspected and where gram-positive pathogens cannot be ruled out.

Again, Dr. Chesney was one of the coauthors of that article, which leads us to the interesting change. We usually had a study drug compared to a single approved agent. In the future, we actually may be seeing combinations of approved agents plus empiric vancomycin for at least the first three days of therapies, for instance, a drug that may have increased activity against a non-susceptible Strep pneumoniae.

[Slide.]

Because of the overall decrease in meningitis in this country, we do anticipate more data from foreign countries, but just a reminder to all of us that the FDA approvals are essentially for the U.S. population, so the pathogens, their susceptibility profiles, and the standard of care of the patient should be comparable to that seen in the United States. That comes as a big issue in dealing especially with Strep pneumoniae and non-susceptible strains.

[Slide.]

Let me start off with enrollment, and you will probably be seeing the same format for every presentation. For enrollment, all efforts should be made to enroll patients with strongly suspected bacterial infections, and we deal with the patients that will be analyzed as a primary efficacy analysis, will be the patients that actually have proven bacterial infection. Therefore, at the time of enrollment and randomization, only patients with strongly suspected bacterial infections should be included.

Thus, the use of a Gram stain result should be strongly considered. This, however, has to be countered with the issue of delaying therapy in patients who are critically ill and therapy needs to be started off prior to Gram's stain results being obtained.

[Slide.]

Why is Gram stain such a big issue? Essentially because inclusion criterion for meningitis are rather hazy. Meningitis is essentially shown if you have a clinical suspicion, and they are essentially based on nonspecific and specific signs and symptoms, so there are problems with that.

First, they are very variable by age. Secondly, in the populations that have the highest rates of

meningitis, especially infants, the signs and symptoms are more nonspecific. Lastly, even if you look at older populations, such as adolescents and adults, where we have all been trained that you see a classic triad, there have been two recent epidemiological studies, the largest actually from Iceland, looking at adolescents and adults with proven bacterial meningitis.

They found the classic triad of fever, headache, and either a positive Kernig's and/or Brudzinski's in about 60 percent of patients, so even the classic triad is less common than expected.

[Slide.]

Exclusion criteria, we have already discussed, and also the potential use of the Gram's stain to not enroll patients that are not strongly suspected to be bacterial.

[Slide.]

Let's discuss the study drug. Because of the severity of this illness, there should be adequate confidence that the agent can get penetration into the CSF. This may be difficult to do and it is probably even hard to convince medical students to do this for $50, so there may be some situations, for example, when you have VP shunt replacements where pharmacokinetic parameters can be evaluated, but there should be some modelling to show that

the drug can get into the CSF and also in-vitro data showing that it is common against the most common pathogens.

On the control drug, concomitant therapy, such as dexamethasone and possibly vancomycin, all three of these should actually be discussed with the agency prior to drug initiation or study initiation, and lastly, at this time, oral relay therapy meaning I.V. therapy in the hospital followed by oral therapy at home is not really standard of care.

That may be with time considering that everything is now done orally at home, that may with time be an actual option, but at this time not commonly done, or actually I am not sure of anybody having even tried doing this, so if this is written, this protocol, this definitely has to be discussed prior to initiation.

[Slide.]

Let's talk about the visits. I will talk about four: entry, on-therapy, the end-of-therapy, and then the two test-of-cure visits which are comprised of the early and the late post-therapy.

[Slide.]

The entry visit. The entry visit, in addition to the full physical examination, there should be an emphasis on a complete neurological examination, coma scale, et

cetera. It is recognized that most of these patients will be entered via the emergency room setting.

There are few care facilities where due to the acuity of the situation and the vast volume in most of these places, it is not common that the physician actually has time to go through a complete neurological examination and documentation. Therefore, they should be reminded in both the case report form and the study protocol that this is very important.

Secondly, the CSF should be sent off for cell count, both red cell and white cell, protein glucose, and a cyto-spun Gram stain in addition to the appropriate cultures.

Lastly, because morbidity rates are actually now part of the efficacy definition, the patient's baseline status of hearing in children, development, and for all patients neurological status should be documented fully for every patient in the case report form.

This does not have to be done initially at the time of study entry, but can be done at a more leisurely pace in the next few days of study therapy and when things have calmed down.

[Slide.]

In terms of on-therapy visit, the only one I am

going to mention is a repeat CSF analysis should be done 24 to 36 hours after initiation of therapy, and it should be a minimum of 24 hours.

Changes in therapy and the addition of concomitant therapy can be done at this time, and usually will be done clinically by the investigator if need be.

Lastly, this is a very good time to get PK parameters of the study drug in the CSF. We had an informal meeting with industry yesterday, and it was brought up by almost every one of the discussion groups that more emphasis should be placed on PK and PD parameters, and since a repeat tap is being done, this would be a great chance to get oodles of information about drug penetration into the CSF.

[Slide.]

The questions will be peppered throughout this talk, and this is actually the first since this deals with the repeat tap, is delayed eradication of Haemophilus influenzae a valid bacteriological outcome for the repeat tap or should this be seen as a failure of therapy?

We will come back to these questions down the road, but I just wanted to bring it up in this proper context.

[Slide.]

The end-of-therapy visit. Most protocols usually

have a range of days of therapy needed per pathogen, for example, Strep pneumoniae, usually between 10 or 14 days, Neisseria 7 to 10 days, et cetera. So, an end-of-therapy visit should be planned for some time in that range period.

The purpose of the end-of-therapy visit is really twofold: one, to see if a continuation of therapy is needed; and, secondly, if a repeat lumbar puncture is needed as well.

For most of the situations, a repeat lumbar puncture will not be indicated. There will be some infections, for example, where this will be the case, such as gram-negative rods, but in the most common situations this will not be indicated.

[Slide.]

Let's talk about the two test of cure visits. The first is the early post-therapy visit, which is to occur approximately five to seven weeks after completion of all therapy. Again, a lumbar puncture repeated only if clinically indicated.

When we get to efficacy definitions, the need to look at morbidity changes in morbidity rates, there should be audiological examination, a developmental assessment, neurological testing done on all patients. Let me do these backwards to keep people awake.

Neurological testing, there should be a full examination, and the findings should be documented. Unfortunately, commonly seen in neurological examination, one box normal, second box abnormal. It would be nice to see actual full documentation of a full exam.

As far as developmental packages, there are several validated developmental packages out there which have been well studied in clinical trials, and one developmental package should be chosen by the sponsor and used in all patients, and this should be chosen prior to the study initiation, then used on all patients.

Lastly, audiological examination, usually, bilateral audio-evoked response test done. In toddlers, it is usually an audiological exam with play or visual stimulation added on to kind of keep them entertained, and lastly, for older children, it is more the traditional hearing test.

The important thing about audiological examination is that an appropriate range of tones should be tested, and there is good literature to show that there is a certain range of tones which are most effect status post meningitis, and those ranges should be the ones that are especially stressed.

[Slide.]

The late post-therapy visit, five to seven months after completion, again, emphasis on hearing, development, and neurological findings. At this time, things such as development of a seizure disorder or behavioral difficulties should be documented, as well.

[Slide.]

In regard to patient population, I know we talked about not having a primary efficacy analysis, so here it is, the next speaker basically doing the total opposite.

The primary efficacy analysis should have a clinical response of the enrolled patients who are what we call fully evaluable, both with a clinical picture consistent with acute bacterial meningitis and bacteriological confirmation.

We may do an analysis of patients who are clinically evaluable only, but the final decision has to be based upon the people with a proven bacteriological meningitis.

So, how do you prove bacterial meningitis? Either having a CSF culture which is positive or in a situation where the culture is negative, but the CSF analysis is consistent with bacterial infection if you have a concomitant blood culture which is positive for a known pathogen.

[Slide.]

More peppering of questions. What role, if any, should the results of antigen testing have in clinical trials?

What is meant here is should antigen testing results be used to enter patients into studies.

[Slide.]

To discuss clinical outcomes, cure is essentially a resolution of all signs and symptoms at the test-of-cure visit and also at both test-of-cure visits, there is normal screening for audiological, namely, hearing, development, and neurologically.

Cure with mild sequelae. Let me start off with the need for predefined parameters for mild deficit. This would really be dependent on the hearings tests done, the developmental package done, et cetera, and it should be written in the protocol prior to study initiation, but this is really resolution of signs and symptoms at the test-of-cure visit early on, so a repeat tap or more therapy was not necessary, and then you had mild deficits noted at the late post-therapy visit.

For people who are missing a late post-therapy visit, then the results of an early post-therapy visit could be used as long as they are mild.

[Slide.]

Clinical failure. I am going to start off with potential clinical failures at the test-of-cure visits and then work back towards the very start of study drug initiation as potential failures that could be carried forward.

If you have persistence of signs and symptoms at the test-of-cure visits, especially necessitating a repeat spinal tap, that should be considered a clinical failure, and also moderate to severe sequelae as defined in the study protocol and again depending on the packages chosen, or the development of a seizure disorder should be considered a clinical failure.

Let's go back to the repeat tap done on the second day of study drug. If there is persistence at that tap which leads to additional therapy or a change in therapy, that patient should be seen as a clinical failure due to the addition of additional therapy. If a pathogen is resistant, however, that patient should be considered unevaluable.

[Slide.]

Going out a little further, at the end-of-therapy visit, if antimicrobial therapy is prolonged for a period sort of out of the ordinary for that pathogen, that patient should be seen as a clinical failure.

Now, to look at the time period between end of therapy and test-of-cure, if you have initiation of either new or further therapy for the treatment of meningitis, that should be considered a clinical failure and carried forward.

Lastly, deaths. Any death that occurs at least after 72 hours of therapy should be considered clinical failure.

[Slide.]

This is the last question. How should patients who die within the first 72 hours of therapy be classified? This deals with the first three days of therapy.

[Slide.]

Microbiologic outcomes. These are mostly presumed responses since the only repeat tap is really done while still on therapy, namely, the second day of therapy. So, the most likely scenario will be presumed eradication, no repeat CSF cultures were obtained after completion of therapy, but the patient was considered a clinical cure.

Documented eradication. For the rare patient who will have a culture done off of therapy, there is no persistence of initial pathogen.

[Slide.]

Presumed persistence. Really two scenarios that we can see commonly occurring, not commonly, but potentially

occurring. One is a change in therapy during the study period, but no repeat CSF culture was obtained. Another presumed persistence is if you have prolongation or initiation of further therapy at the end of therapy or between the end of therapy and test-of-cure visit basically due to lack of clinical improvement, and again no CSF culture was obtained.

Lastly, documented persistence. A repeat CSF culture, and this includes the repeat that we have been talking about, shows persistence of initial pathogen as a documented persistence.

[Slide.]

Let's go back to the questions. First, is delayed eradication of Haemophilus influenzae a valid bacteriologic outcome for the repeat tap or should this be seen as a failure of therapy?

Just to give some background, in the literature there is mention of Haemophilus influenzae persisting after repeat tap and then 24 hours after the repeat tap being negative. Those patients have been called "delayed eradicated." In fact, two of our approved agents have mention of this term in actual labeling.

[Slide.]

The second question is: What role, if any, should

the results of antigen testing have in clinical trials? What we are driving at here is should antigen testing ever be used to enroll patients in trials if there is no positive CSF and/or blood culture.

[Slide.]

Lastly, probably the most controversial one: For people who die within the first 72 hours of therapy, how should those patients be classified? Should we look at all-cause mortality or should we look at these people as unevaluable.

Now, Dr. Chesney will have some comments. Thank you.

DR. CRAIG: Any specific questions for Dr. Rakowsky? Dr. Murray.

DR. MURRAY: I had one. In the document, there is mention that the drug to be tested should achieve or maintain levels equal to or above the expected MIC-90 of the claimed pathogens.

I was just wondering, on the basis of the MIC-90, if you asked me off the cuff, I would have probably said MIC-98, because I think to study a drug in meningitis out, I would be reluctant to study it where 10 percent might be resistant given, you know, with caveats, maybe the resistance is only one step above, and that sort of thing.

DR. RAKOWSKY: That is a good point. I guess I used MIC-90 at that point since it is almost like the lexicon used, but that is a very good point. You have to have a very high confidence in a drug going into a trial.

Any other questions? Okay.

DR. CRAIG: Dr. Chesney.

Committee Presentation

DR. CHESNEY: I would like this afternoon when we talk about acute otitis media, where I have a lot of comments, I didn't really have that much to add to what Alex has said. I guess part of my thought process was that I feel like we have done very well with development of drugs for bacterial meningitis in the past.

Just some comments on his presentation, although I had seen the slides, it is always helpful to have it presented, it raises other issues. I think the first one, Alex, I wonder about is if there should be separate guidelines for neonatal meningitis because it really is such a different entity than acute bacterial meningitis in children, and that was just one thought I wanted to ask you about, because I think the indications for repeat taps are very different.

DR. RAKOWSKY: With these guidelines, we purposely just addressed the most common acute bacterial meningitis

scenarios, and that one slide where we had the exclusion of the following list of patients, in all honesty, separate guidelines should be written for those down the road.

I agree that neonatal meningitis does tend to have a very different, one analysis in two study procedure, for example, there is a high rate of gram-negative rods, there, repeat taps are more common, persistence of gram-negative rods even for four or five days is not uncommonly seen. So, they would probably need their own guidelines, as well. I am in full agreement there.

DR. CHESNEY: Just some thoughts before I heard Alex's presentation, and then I will make some comments based on that.

I think the use of dexamethasone is really a very problematic issue for most of us now. I think there are reasons to think that the mechanism of inflammation in gram-positive meningitis may be different than that for H. flu, and the studies, as you know, that demonstrated a decrease in hearing loss were done almost exclusively for children with H. flu B meningitis, and we just don't see that in this country anymore.

There are no prospective studies demonstrating the same phenomenon for pneumococcal meningitis. In every study there were a few patients with pneumococcal meningitis, and

the best that has been done was a meta-analysis that was published in JAMA a few months ago suggesting that dexamethasone might possibly improve the hearing loss which is substantial associated with pneumococcal meningitis, but I think many centers are not using dexamethasone routinely for meningitis now. I know we are not. We rarely use it in our center.

So, I think any future studies may want to have two arms, one for patients who receive dexamethasone or purposely received it, and the second for those who didn't. Obviously, some of the concerns, certainly if it works to decrease hearing loss, then, that's great and we would all use it, but that is a problem, and the bigger problem is does the dexamethasone alter antimicrobial penetration into the spinal fluid, so that I think any study that incorporates dexamethasone would also have to have a very nice demonstration that the dexamethasone did not alter penetration of the drug into the spinal fluid.

I think the second issue that Alex raised is what now should be the comparator drug. Most of us are very comfortable now with vancomycin and either cefotaxime or ceftriaxone. There have not yet been any failures even with highly resistant cefotaxime strains using that combination. I think it would be important for any new drug, and

certainly we need new drugs.

We are very concerned that vancomycin resistance is going to appear, so we are all very much looking forward maybe to the fluoroquinolones being available to fill that void, and so we want to have new drugs, but whether the new drug could be tested alone against vancomycin plus cefotaxime or whether vancomycin would have to be added, I don't know the answer to that.

I think another big question that is also an issue with acute otitis media is now that the prevalence of penicillin and cefotaxime resistant Strep pneumoniae is so high in many areas in the country, would you need to have a certain number of drug-resistant pneumococcal meningitidis be included as part of the protocol.

I hope I made that clear that rather than just saying that the drug was good against Streptococcus pneumoniae, you would have to show that it was, in fact, good against the drug-resistant pneumococcal cases of meningitis.

The other thoughts I had, I think Alex mentioned obvious penetration of the drug into the spinal fluid, which is clearly critical and we always need that information particularly if dexamethasone is going to be used.

Long-term morbidity, he discussed, and the need

for complete neurologic and hearing testing both at the time of discharge and then again in the follow-up visit.

With respect to your specific questions, is the presence of organisms, specifically H. flu B at 24 to 36 hours, should that be considered persistence or failure, I think to me, and obviously, we very much need the comments of everybody else on the panel, it is rare to see organisms on that 24 to 36 hour tap, but they usually don't grow.

If you had organisms on that tap that grew, to me, that would be a failure, but that is an oversimplified answer to Question No. 1.

The antigen testing, I think it really don't have a place except I could imagine that if you had a positive blood culture and negative spinal fluid culture, but the spinal fluid was positive for the same antigen, but there is such a problem with false positives there that I think that is more of an issue for neonatal meningitis.

I think the question was, is death in the first 72 hours considered a failure, and I don't have a good answer for that, and I would like to hear that from other people.

I think, Alex, those are my comments on your excellent presentation.

Committee Discussion

DR. RAKOWSKY: Thank you for your comments.

I am not sure if we should just put up the questions again and go through them.

DR. CRAIG: Can I get at one of them in a little different way? We essentially have no exclusion criteria outside of anatomical lesions and things like that. What we do tend to do for, let's say, pneumonia, is to use APACHE scores, and if the APACHE scores are so bad we essentially exclude those patients from the study, realizing they were going to probably have a very high mortality, and they are going to be very difficult to evaluate.

Shouldn't we do something like that with the coma score? At least I can speak about this from adults, that you get up to a certain score, your mortality is exceedingly high, and I know that there have been companies with drugs that were unlucky and got all the bad comas on their side, while the comparative agent didn't have any, and as a result, the drug was very slowed in its development.

I think many of those people that have the very bad comas are the ones that died within 72 hours, and so if you can somehow use the coma score to have exclusion of those patients that are likely to have a very high mortality, one might be able to reduce the number of patients that would be dying within 72 hours, and thereby also give a chance of looking at the drug in those patients

where the drug has a good chance of showing a benefit.

DR. MURPHY: It is also the same group that would have sequelae, and I am not sure if you should exclude them or analyze them separately, but the death at 72 hours and the sequelae kind of tend to be the same population.

DR. CRAIG: Any comment, Alex?

DR. RAKOWSKY: Actually, in terms of Botso scores or APACHE's, et cetera, there is some recent literature looking at prism scores in the pediatric population, which is a comparable score to the APACHE where the investigators and actually at Children's here in D.C. have been looking at predictability of prism scores for death in the first five days, which may have some relevance potentially in trying to figure out which patients should not be enrolled.

I was kind of hoping you would bring that up as a potential answer to the third question because it is a difficult issue, I mean for people who die in the first day or two of therapy, you could potentially have, and these studies usually tend not to be very large, 200, 300 patients as a total, so if you have four or five patients who have death in the first 48 hours in one arm, you can potentially skew the results, so it is actually a major issue.

DR. CRAIG: Dr. Chikami.

DR. CHIKAMI: One of the issues in limiting the

population based on a coma scale, for example, and excluding those patients that are more sick is that it limits the patient population, in fact, that you are setting the test drug and its inference of effectiveness that you can then draw, so that, in fact, is a problem.

DR. CRAIG: If that is the case, then, should they be stratified, so that one then stratifies it according to that, so then one is not taking the chance of getting unlucky and getting all of the comas in with your new drug?

DR. CHIKAMI: I think if the feeling is that is an important baseline characteristic, it affects overall prognosis and outcome, then, in fact, that is a consideration in the design of the randomization, and, in fact, that is a characteristic that should be used to stratify the randomization.

DR. CRAIG: Dr. Norden.

DR. NORDEN: I would just like to support that. I think the data from the literature on bacterial meningitis at least in adults and particularly with pneumococcal is that coma is probably the single most important factor in terms of mortality, and so I think it should be a stratifying variable.

In other responses, and then a question for the FDA, I guess, in response to Alex's questions, I think that

I agree with Joan's answer about No. 1, and also if you look at the cerebrospinal fluid parameters, such as white count and if you look at the white count glucose protein, and they are all moving in the right direction, it would make that much easier not to call such a patient a failure.

I think the antigen testing is probably not very useful anymore, and I think we don't do it very often these days. The third question, I think the answer is I think in stratified patients, you have a much better chance of figuring out what to do with patients who die within 72 hours.

The real question I have, that just occurred to me, is the whole question of microbiologic evaluation. I mean basically in meningitis, in pneumonia, bone and joint infections, we really don't get follow-up cultures usually, and so we are always left with presume, presume, presume, and why do we bother with it then in diseases where we are really not--I mean it is just going to mimic the clinical response.

It can't be anything else if you don't have a culture. So, the question is why do we do it.

DR. RAKOWSKY: I guess the response to that would be that what we tend to see is a clearance at the first repeat tap, so in other words, granted there may be a false

negative culture but you are still in therapy, but you would at least have some confidence in the fact that you have one negative culture within the first 24 or 36 hours after initiation, so there is some microbiological data to make a decision on. But in all honesty, most of it is presumed down the road.

DR. CRAIG: I would probably disagree with you, Carl. I think that at least for some of the gram-negative enterics, I think you can see clinical improvement, but not complete elimination where the organism could potentially relapse when stopping therapy, and so in that situation, I think repeat taps are necessary.

DR. NORDEN: No, I am not saying repeat taps aren't necessary, and I agree with you in gram-negative meningitis. I am just saying that most of the time in clinical trials, we don't have bacteriologic evaluation at the end of therapy, and we just call it presumed, so that at least in the trials that I remember participating in, we wound up with just the same numbers in essence for clinical and microbiologic.

I think we go through a lot of contortions sometimes to define bacteriologic responses when they are presumed.

DR. CRAIG: From my reading of the document here,

unless you got a tap two weeks after therapy, it would not be called true eradication, it would still be presumed.

DR. RAKOWSKY: It would be presumed unless you were off of therapy for a reasonable--I wouldn't say two weeks per se--but for a reasonable period of time off of therapy.

DR. CRAIG: Virtually everything we do is going to be presumed eradication.

DR. ALBRECHT: Let me tackle that question a little bit and actually mention a slide that I will have tomorrow in my discussion of the general considerations document. That does involve sort of our use of the term clinically and microbiologically whether evaluable or assessed.

At the risk of repeating myself tomorrow, let me today mention that in looking at these, as you have seen over the past years when we have developed drugs for various indications, we have come into a terminology that we often use called clinically driven indications, clinically and microbiologically driven.

The reason for that is, for example, clinically driven, we have indications where the clinical signs and symptoms identify an entity, and often we will not get cultures, and in the respiratory tract we can readily see

that, you know, there are a lot that are going to be bacterial, but we can also have many that are viral, and, in fact, without a culture, we are assessing patients where we are not positive each and every one whether it was bacterial or viral, so those are the clinical entities.

Then, to the next step, clinically and micro-evaluable, let me by example say otitis. We can have a clinically driven otitis study where all you base the evaluation on is signs and symptoms the child has at presentation and follow-up.

Then, a confirmatory study would be where not only you look at the clinical signs and symptoms, but you perform a tympanocentesis to document the presence of a bacterial organism, and then patients who do not have such a documentation would actually not be included in that analysis looking at clinically and micro-evaluable or micro-assessable patients.

It is true there are very few indications where we actually have the before and after microbiology, and those are what we refer to as microbiologically driven indications, and those would be urinary tract infections, gonorrhea, pharyngitis, and it is really a practical consideration. It is almost routine and possibly we could say trivial to obtain those cultures before and after.

In the ideal world, we would be getting before and after cultures on all patients, but we realize that is not going to be ethical, not feasible, not realistic. So, some of it is the accident of doing clinical trials in patients.

DR. CRAIG: But the point I was trying to bring, I think if you talk to any infectious disease physician, they would consider if the culture is negative after 24 or 36 hours, that the organism has been eradicated, and would not call it presumed eradication, which this document would call it, because it doesn't consider full eradication unless you have one a period of time after therapy has been stopped.

DR. ALBRECHT: Yes, I think that is correct, that you would say if it is off therapy when there is no antimicrobial, that is the definition of the term as we have applied it in the regulation.

DR. MURPHY: So, what you are suggesting is that the discussion of the committee at this point would say, one has a positive culture, meet the entry criteria, you have a 24- or 48-hour negative culture, and patient does not have a follow-up LP, but has met the clinical criteria, that is a cure.

DR. CRAIG: Yes.

DR. MURPHY: No adjective.

DR. CRAIG: That is fine with me.

DR. MURPHY: What we want to hear is your recommendations.

DR. CRAIG: Other members on the committee would agree with that, as well?

DR. MURPHY: Nobody disagrees?

DR. CRAIG: Nobody disagrees.

DR. MURPHY: We have somebody from FDA who would like to comment.

DR. ALTAIE: This is Sousan Altaie. The presumed eradication, because it is the test of cure at 24 hours or re-tap at 24 hours, there is an assumption of antibiotics being onboard and prohibiting the growth of the organism if it was there and if it was supposed to grow.

I think that presumed refers to that and at 24 hours cure to assume cure/cure, we have taken in consideration the presence of the antibiotics in the specimen and the possibility of preventing them from growth, I don't think we can ignore that microbiologically.

We do accept a cure clinically, and at the end when we say we did not have the culture to say it's no longer there, but the patient is doing well, that is why the term presumed comes into play.

I don't think it interferes with the decision of cure at that point.

DR. CRAIG: I agree there is always that hypothetical consideration, but I think there is enough clinical data out there showing that patients that have the organism eliminated at 24 to 36 hours, and have a good outcome, don't have a relapse a week down the line with the return of the organism, so I think you can call it cure.

DR. RAKOWSKY: Maybe to put everybody at ease, when we do a final analysis with a bacteriological outcome, we essentially put together the eradications and presumed eradications as bacterial successes, so I guess I really don't want to spend too much time in terms of arguing semantics of how to call these things since we are going to call them bacteriological successes in the long run. Just to lay that out on the table.

DR. CRAIG: The only reason I bring it up is I think that frequently what gets played in the literature, there is a lot of things that are called eradication that are presumed eradication, and when there is eradication, I think we should give the credit that there is definitely eradication.

Dr. Norden.

DR. NORDEN: To change the subject, one other small qualification, Alex, but in the study drug, I believe one should require that it be bactericidal against presumed

infecting pathogens. I think again there is reasonable data to show that, for example, chloramphenicol, which works against H. flu because it is bactericidal, does not work well against other enteric gram-negatives like E. coli.

DR. RAKOWSKY: That is a point very well taken and will probably get included into the final document.

DR. CRAIG: Did you want to address your other questions? I think the first one was antigen testing. What do people feel about the Haemophilus that is still there at 24 to 36 hours, should that be considered a bacteriologic failure? Dr. Norden thought that if the numbers are going in the right direction, that it shouldn't be.

DR. HENRY: I think that we do want to see the numbers going in the right direction. I would be very uncomfortable if there were still organisms growing at 24 hours. That would be very disconcerting, and I certainly would not want to continue the use of that drug.

So, I think that would be a failure if it was still growing at 24 hours. We were really just talking about the fact that, you know, cultures are negative at 24 hours because the drug is onboard, so if something is growing, that would definitely fall alongside of a failure in my mind.

DR. CRAIG: Any other comments by anybody?

DR. RAKOWSKY: Is it safe to say that the consensus is that it is a failure?

DR. CRAIG: I would call it a failure.

DR. RAKOWSKY: The next question, let's play devil's advocate here. We anticipate more data from foreign sites because of the decrease of meningitis. Another problem in the States is a lot of children are pretreated, so we do have a lot of negative cultures on entry of potential acute bacterial meningitis, so is there any potential way to use antigen testing on more patients in this country? That was one reason for asking this question, as well.

DR. RELLER: Especially for this country, but also for abroad. The persistent inclusion of antigen testing as part of clinical trials, I think should be abandoned. The test lacks sensitivity. I mean the sero-group B would be missed entirely, lacks some specificity, but for purposes of the FDA, the most important thing is that if it does not enable recovery of an organism on which susceptibility testing is based, which is a critical issue here and abroad for pneumococci, and frankly, today, without susceptibility testing and bactericidal testing, as Dr. Norden said, patients should not be included in an evaluation of a drug that is potentially efficacious for bacterial meningitis.

It is time to abandon this test. I mean it got into difficulty in this country with indiscriminate use in a low prevalence population, but for purposes of FDA, there are other, even more compelling reasons to abandon the use of this test for this purpose.

DR. CRAIG: Dr. Henry.

DR. HENRY: I agree with Barth. First, I think it would be hard to standardize which bacterial antigen testing is done, but if it is not going to provide--if you are not left with an organism to test susceptibility, how do you know where to put that patient? You don't know if it's resistant, you don't know if it is moderately susceptible. So, again, I don't think it can be a criteria that you can use.

DR. CRAIG: It probably would have been better when we had more of a single population until all the resistance started to develop.

DR. RELLER: How do you evaluate a patient who has had prior antibiotics and you don't have an organism from them? You don't know whether it was a susceptible one, a resistant one. You don't know whether they really had the disease, they didn't have the disease, whether it's a false positive.

I mean good data, complete data, nothing missing,

with all of the critical elements on a modest number of patients seems to me so much more powerful than missing pieces. You know if important issues are missing, they are worthless.

DR. RAKOWSKY: Is it fair to say the consensus is not to use these?

DR. CRAIG: Yes.

DR. CHESNEY: The one situation where I could see it being helpful is the child and perhaps an adult, and I have seen this situation a number of times, in which the blood culture was causative for pneumococcus, the spinal fluid had pleocytosis and organism, but no organisms on Gram's stain, nothing grew, and frequently the child has a very bad sinusitis, frontal sinusitis, sphenoid, and it looks like a sympathetic meningitis if there is such a thing, and I don't know how you would classify those patients, do they have pneumococcal meningitis or do they have pneumococcal sepsis and sinusitis with a sympathetic meningitis, and if you had a positive antigen test in the spinal fluid there, I think perhaps I would be convinced that it was a pneumococcal meningitis.

I don't know if that is angels on the head of a pin, and I don't know if the rest of you have seen that, but that would be the one situation where I could see that

perhaps the antigen testing would be helpful.

DR. RELLER: Dr. Chesney brings up an important issue. When one has done, in the older literature, doing lumbar punctures on patients with positive blood cultures and particularly those with duration of some time with meningitis, there may be a modest pleocytosis.

In the patient you describe, I would strongly prefer they not be included in the evaluation of a drug for meningitis. As everyone here recognizes, the aberrations in CSF in response to CNS infection, where commonly all of the pathophysiologic work that has been done where the number of organisms is high, why the Gram's stain smear is so sensitive, particularly after sinus centrifugation, those things, even with therapy, do not change that rapidly.

I mean the organism may be gone, but again without an organism, we are stuck, so that these patients who have minimal changes in CSF with a positive blood culture, one would agonize in that situation and let the issue be decided for inclusion or not based on an antigen testing.

I think that these patients are not going to be helpful in the evaluation of a drug, and should be excluded.

DR. RAKOWSKY: Do we have a comment from industry?

DR. GESSER: Richard Gesser, Merck Research Labs.

Actually, I had a question not addressing that

specific question, so maybe you want to wait. I had a question regarding the choice of the timing for the test-of-cure visit for this indication, five to seven weeks as opposed to, for example, two weeks.

How was that decision made and based upon that?

DR. RAKOWSKY: Should we do the question now?

DR. CRAIG: Why don't we finish these up, and then we will do that one then. That was another one I have, too.

DR. RAKOWSKY: This actually is probably the most controversial of the three. I am going to put some context around this. If we see more Strep pneumo than historically the rates of mortality with Strep pneumonia meningitis can range anywhere between 10 and 15 percent in U.S. studies and Central American studies, they range as high as a quarter of the patients, and in those studies, the vast majority tend to die in the first 72 hours. So, this actually does play a role in the rather small meningitis studies that we see. So, I will leave it at that.

DR. CRAIG: Go ahead.

DR. WITTES: It seems to me there are four possibilities. One is if there really is an unstratified recruitment and people are coming in, and everybody is entered, I think you have to include them, the 72-hour deaths.

The second possibility I see is actually doing an exclusion from the trial on the basis of either the coma score or some very high-risk variable or set of variables, and then the issue is moot, because they are not there.

The third thing would be actually similar to what you are suggesting, stratified, but think of it as two protocols. One was just a protocol that you are looking at the "low risk" group, and then another protocol there is just a small strata of high-risk people that you are sort of looking at in an exploratory way, is this new drug going to do anything for these people, but not up-front excluding them from the analysis.

The final possibility is to stratify on the basis of this high-risk strata and then analyze by strata, and that should, in fact, reduce the variability, but then you are including the deaths, but comparing them within the strata in which they have entered.

DR. CRAIG: Sort of like your No. 3.

DR. BLACKWELDER: I agree with Janet on including them as failures, but there is more than one way that might be valuable to look at the data, and it seems to me--I am assuming you are going to exclude them because you feel the drug could not have prevented that.

DR. RAKOWSKY: I am not going to reveal my hand of

cards here.

DR. BLACKWELDER: I will make that assumption. For whatever reason it seems to me that if that is a possibility, that you might want to do an analysis without these deaths because including them could falsely lead you to the conclusion that two are very similar or equivalent.

Finally, a naive question. Isn't it at least possible that you would want to look at that event separately, because one drug might be able to prevent these early deaths, and another one might not?

DR. CRAIG: Yes. My feeling, I just like stratifying at least according to coma because I think that does have a major effect on outcome, but obviously, that is going to be a smaller percentage of the patients. You are probably not going to get enough of them to really do much statistics, so I agree it is sort of a group to sort of see if the new drug is going to be wonderful and even do something in there that wasn't seen before, but in terms of the other group, I would consider these patients--I assume they are going to be equally divided--I guess I would have to consider them failures if they didn't have severe coma and something like that in the beginning.

Now, I know where he is coming from because there is the old study with pneumococcal bacteremic pneumonia,

that there are always about 5 to 10 percent are deaths, and again the question is are these patients with meningitis also the bacteremic ones, and are we just seeing this death that has been well described to occur in a significant number of patients within the first two to three days, and that is something that you are going to expect to see.

Well, I think if you got numbers, I would expect you to see it in both groups if you sort of factored out the problem with severe coma, so I would call them failures, but I expect that you would find them with the comparative agent, as well. If you don't call them failures, you are not going to give the new drug a chance to even work in those people.

DR. CHESNEY: Could you not separate out microbiologic failure from clinical failure? Many of the children that we see who die very abruptly have very susceptible organisms, and the spinal fluid is sterilized very promptly, but they die obviously for other reasons that we are talking about, so would it not be possible to include them as microbiologic success, but clinical failure, or call it something else?

DR. MURPHY: Joan, they would end up as a failure. I mean I think the important issue here is that death could tell us something, and we should find a way of evaluating

it, and I think it optimizes the information that we can get from the study of the drugs.

DR. GOLDBERGER: I actually have not been previously involved in meningitis to date, and I certainly would agree with the issues of attempting to avoid balances between treatment arms and actually the opportunity perhaps to look separately at the more severely ill patients.

Having said that, it seems to me that there is no option other than to consider these as failures. I think your point is a good one, but there will be a group of people who may be predestined to die very early on. There is, however, no reason to believe that if you have a somewhat less effective therapy on top of that group you can't add some other patients, as well.

DR. CRAIG: So, your three questions. Could we then have the question from the audience?

DR. GESSER: The question was regarding timing of a test-of-cure visit, why five to seven weeks as opposed to, for example, two weeks?

DR. RAKOWSKY: Actually a good question. In the '92 Points to Consider, the timing was not actually placed in there, but the IDSA-FDA guidelines from approximately the same time actually mentioned the five, seven weeks, and they actually have a six- to eight-month time period.

That was based on the literature at the time that those guidelines were written in terms of having at least some appropriate post-inflammatory time where the patient can actually have some decrease in swelling, et cetera, so they chose five weeks as a minimum--again, I was not involved with this--but that was the designation by IDSA guidelines at that time as a way to see a child who is more back to normal status after a meningitis, and then a late follow-up, they called it six months, and we just went with the same numbers of five to seven months, and if anybody wants to add historical data who has been here longer than I have.

DR. GESSER: I guess in the interest of getting complete data, i.e., full follow-up on all patients, certainly if there is more information that can be gleaned in those three to five weeks that are additionally added with five- to seven-week follow-up, then, it is certainly something to go after, but if that is questionable, perhaps in the interest of assuring follow-up in more patients, maybe a shorter time period should be looked at.

DR. CRAIG: Dr. Chesney.

DR. CHESNEY: Alex, I don't know where that came from, but I know that Ralph Feigen did a very large follow-up study of children with meningitis, and it was

remarkable how many who were either quadriplegic or paraplegic or stone deaf at the time of discharge, within--and I would have to go back and look at the paper for the time period--but within several months afterwards, some of them were up running around who had been paraplegic, and the hearing was restored in a number of children when you waited several weeks or months to look at them again, so that may be why that time period was chosen.

DR. ALBRECHT: Again, I don't have the answer to why that was, but from a historical perspective I think, unlike many other indications for meningitis, the follow-up in the older days, if you will, tended to be around a month, and it was I think a compromise trying to capture both the bacterial response to the drug, as well as sort of the preliminary neurologic sequelae that may or may not have resulted as a consequence of the infection.

I think now clearly, based on the IDSA guidelines and our current thinking, we look for the neurologic sequelae at a later time point, the five to seven months, so certainly I think it is worth discussing whether that early visit should be two weeks, five to seven weeks, four weeks, or whatever, but realizing that really we are equally interested in that long-term follow-up, and occasionally we have had submissions that have looked at the patient at one

month after completion of therapy, and did not have a later follow-up, and it is extremely difficult to assess the long-term sequelae in such cases.

DR. CRAIG: Is it possible to have--I mean do you have to have separate times? Could you have a two-week one, and then if there is neurologic sequelae that they have one later on, so that you can try and make sure you get data? As I said, the longer it goes on, the greater chance of losing follow-up with the patient.

At least if you did it early, and if there was neurological problems, to do it later, then requiring everybody to do it that far away, that might be helpful for the companies to collect the data.

DR. RAKOWSKY: I guess an ideal situation--I know it's a very good question--an ideal situation if you do have consistent late follow-ups, then, the early follow-up can be more variable, but again, if you are trying to capture patients who do not have a false positive sequelae, again, Feigen's article comes to mind. I think it was six to eight weeks in that article. After that point, the patients who appeared to have some permanent sequelae appeared to also have resolution of them. So, you want to have a long enough episode of time afterwards.

DR. CRAIG: But if a person didn't have any

neurologic sequelae, we are waiting five to seven weeks before we do any follow-up on them. Do we need to wait that long and take a chance of that patient being lost to follow-up?

DR. ALBRECHT: Well, could I maybe ask to say that, to rephrase that as two questions that the committee could perhaps give us some comments on. One is when would that first visit be appropriate, and then depending on what is found at that first visit, you know, should we actually be saying maybe the second visit isn't critical in certain subsets, and is only necessary in patients where something is found at the first visit.

So, I wonder, I think clearly the recommendation we are making right now is we would like an early visit and we would like a late visit for confirming what the neurologic sequelae are if they exist. But should we be rethinking that? Could you all comment on how many visits are really necessary, and on what criteria those visits should be based.

DR. CHESNEY: I think the first complete evaluation should be at the time of discharge, and I think the second one, I think Bill makes a very good point, if you have no neurologic deficit at the time of discharge, but that is often hard to evaluate because for younger children,

they have been in the hospital and they may not be relating well, and what have you.

So, to me, I think this is certainly something everybody should think about, I would probably suggest one more visit maybe a week after discharge or two weeks after discharge when they have been home, they have been off antibiotics, they are back to their normal lifestyle, and if that visit they are totally normal biologically, then, I agree with Bill, I don't think there is a need for a visit after that, that that should be restricted to the children that clearly have some demonstrable developmental or neurologic impairment at that second post-discharge visit.

DR. CRAIG: What percentage in a clinical trial would you say would be relatively normal at two weeks, let's say one to two weeks post-therapy?

DR. CHESNEY: Well, since we are dealing with pneumococcal meningitis now, for the most part, and I think the figure is between 20 and 25 percent have some hearing impairment, it is going to be at least 20 to 25 percent are not going to be completely normal.

DR. CRAIG: But 75 percent would. I mean I think the companies always have to look at the cost of multiple visits and trying to keep tracking, but if you could knock off 75 percent of them, you are only down to following 25

percent for an extra time, and taking the chance also of not losing them for follow-up. I would think it would be worthwhile to do one at one to two weeks after therapy.

DR. CHESNEY: I think the 25 percent is a conservative estimate because there will be another group that have other neurologic changes, but I would think probably a minimum of 50 percent would be well at the post-discharge visit.

DR. GESSER: I think you are assessing two things, you want to assess two things. One is bacterial clearance in the absence of recrudescence or recurrence, which you can assess in the shorter time frame, and the other, correct me if I am wrong, but the intent is to look at perhaps there is a difference in the sequela, which is a more long-term follow-up, but if the test-of-cure is really the bacterial response and lack of recrudescence or occurrence, then, I think you can assess that. I would suggest that you might be able to assess that at an earlier time point than five to seven weeks.

DR. RAKOWSKY: Actually, the test-of-cure would be a combination of both microbiological and clinical response plus the effect on morbidity. I am referring back to neonatal ischemic models where you can actually have children four to six weeks after an ischemic attack

post-birth appear to be normal neurologically, and then develop cerebral palsy several months after discharge.

So, we really have to look at the literature in regards to how long it takes to actually see, and that would have to be more, at least with the hypoxic ischemic event literature, there appears to be a time lag in terms of actually seeing developmental changes or motor score changes, and since a lot of these patients will be ranging towards that age group, that may be something we could use to model.

DR. CRAIG: I would agree with you. If there are problems that develop after therapy and where there is a significant delay, and it is not one in a million occurrence, then, I can see the reason for pushing it all the way out, but if the primary reason for going the longer period of time is to allow something that is present to resolve, I am not sure that one needs therefore to use the very long time period for all the patients.

DR. RAKOWSKY: But again if we are looking at a model where you potentially won't see any problems at two weeks, yet see developmental problems five months down the road.

DR. CRAIG: I said if there is data for that, and it is not one in a million, where it is a very infrequent

event, then, I would say yes, there is justification then for going out for the longer time period, but if that is not the case, and the incidence of what you talk about is exceedingly rare, and with the number of patients that you are having, it is not likely to occur except just by chance, then, I don't think that one needs to go to that extreme, but I am willing to see what the literature says.

I think it is something you need to look at. I think what we are trying to say if that risk isn't there, and it's not a significant risk, that looking at an earlier time period would be appropriate for those that do not have neurologic sequelae.

DR. RAKOWSKY: Maybe another way to look at this, to respond to the two points of the test-of-cure, if we are looking at bacteriological and clinical confirmation of resolution, maybe a two-week time period will be more reasonable that a five- to seven-week, and then to look at sequelae, to do a late visit at that time. Again, we need to do a literature search to figure that out.

DR. CRAIG: Yes, Dr. Henry.

DR. HENRY: I think that you really have better compliance with follow-up visits. Most parents, at least looking with the pediatric population, most parents will be glad to come back and make certain that you can assess their

child and tell them that things are going fine, so you will gather more data because if you tell them, you know, you don't need to come back until a later time point, you may lose a lot of people to follow-up. Again, if you don't have the data points, missing data is, you know, no patient, no data, it doesn't give you anything towards the study.

Again, if you find out at that two-week, or 10- to 14-day time frame, you see the patient that there is something wrong, again, compliance is going to be better because those parents will come back and want to make certain that they know the status of what is going on at the time.

DR. MURPHY: Could I summarize what I am hearing then?

DR. CRAIG: Yes.

DR. MURPHY: Basically, what the committee has said is they feel that we need an acute follow-up to assess that the acute process is truly resolved and then assuming that it is an issue--and we do think it is for long term--that we need a long-term sequelae follow-up. Is that accurate?

DR. CRAIG: I may need to word it a little different. What I am trying to say is that a long follow-up may not be necessary in all patients and that the reason for

doing it is to not lose patients to follow-up. I am not saying that we feel that there has to be one done at two weeks. Sure, you can do it out at five to seven weeks, but the problem I think we are saying is that you may lose some patients to follow-up by doing it out that far for everybody, and there may be some patients that can be evaluated at an earlier time frame.

I don't want to tell the companies that they have to do an entirely new follow-up when I think we would all say that the five to seven weeks would still be okay, but for some people you might lose them, so, we are saying that maybe an early one may be beneficial.

DR. RELLER: I was not present at the discussion of the five to seven weeks, but it seems an odd time and more like it was a compromise between trying to do both the early and the late follow-ups at one visit, some middle ground.

The late visit, just to clarify, and I don't think it has been fully resolved in the discussion, there are two possibilities for the late visit. One is the resolution of things that looked awful, but got better, which would be a great relief for confirmation for parents, for example, and in accord with Dr. Henry's comments.

The other is that may not be there with the

literature review of something that was not seen, so something early on that resolves late, and something that was not present early on that appears late, and if the data are not there for that being a reasonably frequent occurrence, you would skip that for the majority of patients.

So, why not have an early visit, like two weeks, and then for those who are abnormal, unless the data suggest that surprises come up, a time consonant with the literature that would encompass a reasonable time for resolution, if there is going to be resolution, maybe that would be six months, so two weeks and six months, that is what data support to really define why one wants to see the patients at those two times, and the latter one probably being a subset, 25 to 50 percent of patients who had demonstrated abnormalities that one wanted to assess resolution six months down the line.

DR. RAKOWSKY: I guess we are dealing with three scenarios here. With the first, you have a patient who is normal at the follow-up. Let's assume that both bacteriologically and clinically the patient is doing fine at two weeks, we have answered one part of the test-of-cure equation.

The second part is looking at the development,

behavioral, et cetera. You can have a patient who is normal at two weeks and normal at a late follow-up, which hopefully will be the majority of patients.

You can have a patient who is abnormal at two weeks and then due to inflammatory changes resolving, et cetera, a normal exam or just mild sequelae at six months, so you can take a clinical failure will then become a cure.

I guess our fear would then be also that you have a patient who is normal at two weeks, but then develops problems at six months, and that is where the literature search has to come in.

One potential is two years ago there was a meta-analysis, and there are problems, of course, with meta-analysis of viral meningitis in this country where behavioral issues weren't really seen until six months, children were coming back at a monthly basis, and they followed approximately 300 children, and actually behavioral issues, as well being followed, so it's a different, you know, potential discussion plus meta-analysis, but again there is that fear that you may have a normal child at two weeks, and yet you may miss sequelae down the road at six months, so we have to make sure we capture that third population, as well.

I mean it looks like we are capturing the first

two populations, that give a normal and they stay normal, if we don't have late follow-up we capture those. If we have an abnormal, they don't necessarily come back for six months, you capture those, but the one group that we may miss by not mandating a late follow-up is the normals at the two weeks and then potential abnormals at six months.

DR. CRAIG: I agree that you may miss them, and I agree that you need to look at the literature to see how common that is.

DR. CASE: Gludi Case, Bristol Myers Squibb. I just want to make more general comments about the studies to be done in this indication, and maybe on a similar point that Dr. Goldberger asked us this morning about the delta and change being indication specific.

As was indicated in the presentation from the FDA, this indication, meningitis is fairly difficult to study in the U.S. for many, many reasons, and the availability of patients is difficult. Many sponsors, many of our companies do studies outside of the U.S., mainly Latin America and Central America.

In those countries, availability of disease is somewhat more severe than what we see here because of later interventions, and the early intervention is one of the reasons why these studies are difficult in the U.S.

At the same time, we will anticipate to have cure rate as defined here to be relatively low, and we heard, for example, 25 percent sequelae in patients with Staph pneumo.

If we have to apply the principle that we heard this morning of having a delta of 5 percent for this type of indication, and the cure rate in the 70 percent are eventually lower, we will have further the difficulty of conducting these trials to a level which may make them almost impossible to have.

We know that these are indications which are difficult to recruit patients, where we have to go outside of the U.S. and recruit patients with fairly severe disease and potentially different disease than what we have in the U.S., and at the same time we are facing in terms of demonstrating efficacy, and that will make these trials even more difficult to perform.

DR. RAKOWSKY: That is a very good point. Let me start this response back, and then Paul or Daphne, or whoever else. In regard to the studies, I mean one of the major reasons for a double-blind comparative study is that you have a closed scenario. We potentially will be seeing higher morbidity and mortality rates than in prior studies done in this country because of potentially greater amounts of Strep pneumoniae, and secondly, because of the different

quality of patient care in some of the foreign countries.

So, if we have a well-balanced study, well-blinded study, then, the comparison would be not so much the historical perspective, but in that study per se. I will leave the statistical considerations of then dealing with the lower efficacy rates to whoever is up to the task. But it does make it more difficult.

I guess one reason for mentioning that for meningitis is that the confidence of having a drug equivalent to a proven drug on the market already, which appears to work for meningitis, is one of those indications where you do not want to be wrong as the medical officer of approving a drug, because of the potential complications of patients who do fail.

So, I agree with a stricter comparison. It may be more difficult, but we have had one recent approval using stricter comparisons, and again done in areas of the world where mortality rates are high, and yet it fell into that range.

Further comments?

DR. MURPHY: I think the point is well taken that you will have the comparator drug in the same circumstances.

DR. CRAIG: My question comes up is let's say you are comparing it with a single agent, and you pick one of

the cephalosporins for which we do have resistant strains, will those cases essentially be discarded then?

It is one of the things that we have always had with concerns with the industry is how do they get their drugs approved against resistant organisms if resistance excludes them from the clinical trial, or should they all use vancomycin along with the cephalosporin for pneumococci in order to ensure that they have a comparative agent that should give very good results even for resistant strains.

DR. RAKOWSKY: There are really two ways to tackle that and it depends on the future labeling of the product. If you have a sponsor who develops a product with very good gram-positive activity or activity against penicillin and cephalosporin nonsusceptible strains, and they have the confidence that this drug will be equal to a combination of vanco and ceftriaxone, then, you can potentially have a study done and labeling done where the resistant strains are then included into that claim, because you have essentially shown that you are active against those strains as a single agent.

The other scenario is if you have an agent which does not have that activity, let's say there is another third generation cephalosporin that comes along, looking for a meningitis claim, then, the question comes up do you have

to include patients with nonsusceptible Strep pneumo in that study.

From a clinician viewpoint, you want to have some security that it covers those strains, but from a regulatory viewpoint if the labeling sought is for pen-sensitive and cephalosporin-sensitive strains of Strep pneumo only, I guess I wouldn't have a problem in terms of having a study where they would either exclude those strains or use vancomycin empirically as long as the labeling somehow would mention it.

So it really depends on what the potential labeling of that agent will be down the road.

DR. CRAIG: But aren't the meningitis studies a little different than many of the other kind of studies in that we do get bacteriologic assessment within 24 to 36 hours, oftentimes before you know what the susceptibility is on the organism, so that really you can sort of have everything up-front, and if there is still organisms that grow out of that second culture, we are calling them bacteriologic failures, and if they are negative there, even though it later turns out to be an organism that is resistant to one of your potential agents, you have already got a value back that you are at least having either failure or bacteriologic cure.

So, isn't it okay to up-front, I mean I think with meningitis, what I am trying to say is I think you don't have to toss out anything, you can sort of take it all up-front, because you are going to get within 24 to 36 hours, you are actually going to get a bacteriologic test of whether the drug works or not, and that is going to be before you get back your susceptibility results, because the susceptibility results are usually not back really earliest on day two and oftentimes day three.

DR. RAKOWSKY: Let me answer that by playing some devil's advocate here. Let's say we have a scenario where cefotaxime is approved, and ceftriaxone is a novel agent being studied for meningitis in the present day scenario of strains.

Cefotaxime gets vancomycin added to it because of potential nonsusceptible strains of Strep pneumo. Ceftriaxone, because of similar MIC profiles, the cefotaxime would also have vancomycin added to it empirically in this study.

In that kind of drug study scenario, where you have concomitant therapy added on because of potential resistant strain, I guess the question comes up will we not approve ceftriaxone because it does not cover the nonsusceptible strains, yet, it is as good as an approved

agent for the susceptible strains. That is one way to look at the study.

DR. CRAIG: I see what you mean.

DR. RAKOWSKY: Another way to look at the study is to say ceftriaxone by itself will be used in this indication, and they claim that ceftriaxone can cover as well as the combination of cefotaxime and vancomycin, and that arm would be considered a failure because the claim was that they could cover those strains, as well.

So, looking back at the first scenario, it would be almost unfair not to approve that agent if the labeling states specifically that this agent is not approved for nonsusceptible strains of Strep pneumo and if it is clearly done and the advertising is clearly done.

We hope to get drugs out there that could be used as sole therapy for all strains at this time, but on the other hand, would it be fair to not study agents that are as good as the ones we have out there at the moment as long as you have the proper labeling caveats.

DR. CRAIG: Dr. Reller.

DR. RELLER: Alex, in this situation, before considering studying a new drug, in the picture you have painted, wouldn't you have to have already shown for the resistant ones, that the drug was comparable, or a separate

or preliminary study before the ones that were potentially resistant or turn out to be resistant, that you have shown comparable activity in the susceptible strains?

I mean I could envision--I mean the example you gave with cefotaxime and ceftriaxone was a great one, but what if you put together a new compound that was actually inferior to one of those, with vancomycin, with pneumococci, and it didn't look any different from ceftriaxone and vancomycin because of what vancomycin was adding, and you actually got a drug that was approved that was inferior.

DR. MURRAY: [Off mike]

DR. RELLER: But then you have the vagaries of that. I mean I would hate to see the companion drug to vancomycin being a lesser drug that somehow was able, by default, to get through the process. You would be totally dependent upon the vancomycin component for the comparable, similar efficacy of the combination compound. Actually, it will be dicey situation it seems to me.

DR. CRAIG: Ethically, probably the only place you could study it, otherwise, it would be in a place where there is essentially no resistance.

DR. RELLER: What I am coming around to is, is it reasonable that one could at this juncture only study for a pneumococci in these drugs that would be based on

pharmacodynamics, in-vitro activity, safety. I mean the early trials, something that would be alone.

It is very difficult for me to imagine how you could couple a drug with vancomycin and study it against ceftriaxone or cefotaxime and vancomycin, and come to a reasonable conclusion unless you pitted a drug alone against that combination.

DR. RAKOWSKY: What may be saving us is just the dynamics of marketing. I mean drugs were developed essentially because of an advantage. To our advantage, at this time, we have two agents out there that are both safe and effective.

Theoretically, you can have an agent which is equal to those two cephalosporins being developed for, let's say, a dosing interval, but you already have a q24 or for a safety reason but you already have two safe drugs, so what we envision under those agents that have some advantage of those two, which namely would be probably better activity against nonsusceptible strains.

I guess this would be much more a problem if we are dealing with choices where you have q6 dosing, q4 dosing, and fairly toxic drugs, but in this situation we are in, most companies probably wouldn't--and I will leave that for the companies to respond back to--wouldn't go developing

new agents if you already have very strict competition if you don't have any advantage over those two.

DR. GOLDBERGER: I was actually going to ask Dr. Norden a question which maybe has been broadened a little bit by some of this most recent discussion. I was originally going to ask, when you talked about demonstrating bactericidal activity, the test tube and the spinal fluid are very different. I was going to ask you if you wanted to elaborate on some of the models or some of the other factors you might particularly want to accentuate.

Listening to this most recent discussion, the question then comes up more broadly about how much information we would like about a new compound before it goes into a larger scale clinical trial, which may be whether preclinical data including possibly animal models is sufficient or whether there needs to be some pilot studies in humans, as well.

I don't know if you or other committee members want to address that.

DR. NORDEN: Well, I will start. I think you can do bactericidal activity in a test tube, I don't think you have to do it--without having spinal fluid there. I mean that is an issue.

I think my reading of the animal models--and I am

certainly not a meningitis model expert--is that they seem to me to be fairly predictive of what happens in the clinical situation, but I would defer to somebody like Bill.

DR. GOLDBERGER: I guess my question about the test tube was depending on the class of antimicrobial we are studying, there may or may not be a significant inoculum effect which you can deal with in test tube. There may or may not be a significant pH effect.

Whether or not people would on their own necessarily cover all those issues or whether there needs to be special attention because of the kind of model we are using in therapy, that is what I wanted to know whether you want to elaborate on.

DR. NORDEN: I don't have any information to really add to that.

DR. CRAIG: I agree. I would tend to think that the animal models are fairly predictive, and nowadays where one can actually simulate human pharmacokinetics in animals, that one can do it even more realistically than what has been done in the past where we have been more dependent on animal pharmacokinetics, which obviously are different than what we see in humans.

I find it hard to think of a situation, but I mean I think the problem we have with animal models is we tend to

take one organism and generate to huge populations, so I think it is important that one look at a variety of different organisms, some resistant ones, as well as standard susceptible ones, so that one can see how the drug behaves for a variety of organisms, and not just basing it on one single strain.

DR. RAKOWSKY: I guess just a question of naivete here. The protein response in animal models, would that be as great as you see in humans? I guess I am driving at the potential protein binding of an agent, would you see that in animal models, as well.

DR. CRAIG: I think when you start looking at the levels and everything that you get with many of the drugs that you are talking about, that are highly protein bound, where you have titers that are like a 1,000-fold over the MIC, or 500, you have to have a heck of a lot of protein binding in order to really significantly reduce that.

I mean I think there is some theory that you can look at ahead of time to bring in that question, and therefore use a dose that might be at the lower end of what someone might see, so that you can then probe that with the organism to see if that does come out to be a significant factor.

So, there are a variety of things that you can do,

but I think you can get a lot of these answers in animal models, and that you don't have to necessarily do anything more than getting some kinetics of the drug in humans before embarking on clinical trials.

Dr. Chesney.

DR. CHESNEY: I was just going to confirm that. I think our use of vancomycin and cefotaxime has been based almost exclusively on George McCracken's rabbit model, and he was able there to use the dexamethasone and highly ceftriaxone resistant strains, and to show that you could get perhaps synergy in that setting, and combined with the in-vitro data is why we started using them together, was totally based on the rabbit model, but it turned out to be a very good predictor.

DR. MURRAY: With regard to your question about inoculum, I think if you are looking at endocarditis or at meningitis, you ought to do the cidal activity of both 10^5 and 10^7.

DR. CRAIG: Any other comments? The FDA feels they have enough input back on this particular topic? Okay.

DR. RAKOWSKY: I would just like to thank Dr. Chesney.

DR. CRAIG: We will now have our lunch break. We will start at five after 1:00.

ajh

[Whereupon, at 12:05 p.m., the proceedings were recessed, to be resumed at 1:05 p.m.]

A F T E R N O O N S E S S I O N

[1:15 p.m.]

DR. CRAIG: We can move on to acute otitis media. The FDA presentation is by Brad Leissa.

Acute Otitis Media - FDA Presentation

DR. LEISSA: My objective for this afternoon, it is a little different from what we just went through with bacterial meningitis. Bacterial meningitis, this morning was the first time that that indication was presented to the advisory committee as an indication. We went through this indication back in March of '97. This is why we are at round two.

What I am going to attempt to do over the next half-hour or so is to remind, recall for people on the advisory committee, as well as in the audience, what discussion we went through a year ago, also recognizing that there are people on the advisory committee who were not part of the discussion a year ago, and therefore, of course, are welcome to opening any discussion to issues that I may not actually be bringing up that they may be reading into the document, but they also believe needs to be addressed further.

[Slide.]

What I would like to do is remind people of the

questions that were posed a year ago, the responses that came to those questions.

The first question that was asked, again back in March of '97, the issue was for the clinical-only--and this was the idea that there is in otitis media, there typically has been a clinical-only study and a clinical-microbiologic study, two different studies--but for the purposes of the clinical-only, acute otitis media study, in the interest of increasing diagnostic specificity at entry, should the guidance recommend minimal baseline clinical findings and/or tests for evaluability, for example, tympanometry or electroacoustic reflectometry where age appropriate.

The issue here is the concern that we had in the division was that in this clinical-only study, were we actually including into the study population a number of children who did not actually have a bacterial otitis media, but that there were other viral causes or other nonspecific causes that would typically present clinically the same way, therefore, the issue of differentiating out a bacterial versus a viral versus other presenting signs and symptoms.

[Slide.]

So, what came back from the advisory committee, again a year ago, and I have summarized some of these paraphrased, is there was consensus about wanting to

optimize specificity and to minimize inter-investigator variability.

In doing that, one of the recommendations was to standardize otoscopy amongst the investigators, and specifically, one of the consultants recommended biphasic pneumatic otoscopy. The issue there would be to verify the TM initially through insufflation and then to exsufflate to see mobility of the TM. That is what is meant in the consultant's perspective of biphasic otoscopy.

Secondly, "Tympanometry and electroacoustic reflectometry are practical and should be required."

Another comment was, "We need a bulging tympanic membrane on study entry."

[Slide.]

So, in keeping with the feedback that we received from the advisory committee members is that we have done in the revised guidance document is that we have strongly recommended at study entry, in the interest of specificity, for patients with bacterial infection in that clinical-only trial especially is the presence of a bulging tympanic membrane. Also, on top of that would be the biphasic pneumatic otoscopy consistent with a middle ear effusion, and tympanometry or electroacoustic reflectometry consistent with effusion, as well.

The caveats that I have added in here is that at some point it may be "unethical" in the presence of a bulging TM, to insist that these be done and that in the presence of a bulging TM, that may be enough clinical information to go with.

[Slide.]

The second question that was posed to the advisory committee: What is the appropriate timing for the acute otitis media test-of-cure visit independent of the pharmacokinetics and pharmacodynamics of the drug? Specifically, is a one to two weeks post-therapy sufficient time to assess a drug's efficacy in the treatment of otitis media? A similar question that was raised this morning about with meningitis about what would be a truly practical and an optimal time for the test-of-cure visit.

[Slide.]

The responses that we got back from the advisory committee were the test-of-cure visit at one to three weeks post-therapy may be most reasonable.

But then there was another side making the recommendation of a three- to five-day on-therapy visit is very important, as well, as in these studies.

[Slide.]

So, what has happened to the revised guidance?

What we have recommended is the test-of-cure visit occurring two to four weeks after entry into the study. It is consistent in the scope with regards to the comment of the one to three weeks, but the idea about two to four weeks after entry into the study is recognizing that with future trials we may see shorter and shorter durations of therapy where the comparator may be at a different timing, and therefore it seems to us the most reasonable to be using the same test-of-cure visit relative to time at entry, especially if you are saying, for example, that a 10-day therapy is similar in efficacy to a five or seven day, that in doing that, that that same standard of test should be applied, which is relative to after entry into the study.

[Slide.]

The third question that we posed: In light of the 1992 Divisional Points to Consider document, for the clinical/microbiologic study, is the evaluation of 25 Strep pneumoniae sufficient, depending on the drug, in light of increasing concerns about resistance or should greater Strep pneumoniae experience be sought in designing clinical trials?

[Slide.]

The response that we got back on this was varied. There was consensus that pneumococcus is our biggest

problem, but one person said that we need as many as 100 Strep pneumoniae to be able to say anything in this very important pathogen.

[Slide.]

What we have said in the guidance specifically is trying to back away from this finite number of 25, because I don't think we have the answer to that, but we have stated in the guidance that Strep pneumoniae resistance has become an increasing concern for the global medical community. Strep pneumoniae is the major pathogen in acute otitis media.

Because of this concern, 25 patients with Strep pneumoniae may be insufficient to garner approval for this pathogen in acute otitis media. Greater certainty in the investigational drug's purported efficacy against this pathogen may be desirable.

From the sponsor's perspective, I am sure they would be thinking I wish we had a number, can you be any more specific than that, and I invite the advisory committee to give us any more guidance, and we also, of course, need to hear from industry what is practical, what is doable.

[Slide.]

Question 4. Depending on the drug--this was again a year ago this was asked--depending on the drug, for

example, for beta-lactam, should acute otitis media clinical studies be conducted in geographic areas where Strep pneumoniae resistance and/or beta-lactamase resistance are known problems?

[Slide.]

The response that we received, "Geographic location is becoming less of an issue as resistance increases everywhere. However, because of this problem, we need more tympanocenteses to address efficacy in resistant pathogens."

Another comment, which is a little different, because the first one implied, well, it is not such a big deal where you do it, but we just need the tympanocenteses, the second comment was, "We need patients enrolled from across the U.S. including areas of high-risk resistance."

[Slide.]

What have we done to the guidance? Not feeling that we got very clear guidance to go one direction or to change what we had had previously, haven't actually changed the guidance, we have not revised it, so the issue of recruiting study centers by geographic area--where increased drug resistance prevalence may exist--is not currently addressed in the document.

Again, we invite the advisory committee to guide

us further on this if you all believe that that should be changed.

[Slide.]

Question No. 5. The IDSA/FDA guidelines--again, these are the guidelines that were published in Clinical Infectious Diseases in 1992--suggest: "Patients should be followed up clinically and by otoscopy biweekly until middle ear effusion has completely resolved. The time to resolution of middle ear effusion should be recorded."

[Slide.]

Reading the IDSA guidelines, this implies that a drug's efficacy claim for the treatment of otitis media should be linked to middle ear effusion resolution? The issue is, should otitis media clinical trials, for the purposes of regulatory drug approval, be designed to assess the time to middle ear effusion resolution?

[Slide.]

After some discussion--again, this was a question that was posed a year ago--the consensus that we got from that was something of interest, the issue of middle ear effusion, something to be studied, but relative to an evaluability criterion, where you have patients required to follow-up for monitoring resolution of ear effusion, this is not what we are looking for.

[Slide.]

What change have we done to the guidance? We have added into an Analysis Section--which I will talk about a little bit later--but the idea of actually looking at this and making that as a secondary efficacy analysis.

[Slide.]

Those were the actual former questions that we posed to the advisory committee, but obviously some more discussion came of that, and this was actually a question that I believe Dr. Reller brought up at one point, the issue about the clinical-only study, whether it is germane to demonstrating effectiveness in otitis media.

One comment that came from that discussion from one of the pediatricians was, "I can probably think of 10 people, given enough time, who can do tympanocenteses. I would have to walk in the manufacturer's shoes to know whether this (requiring tympanocenteses for all patients) is feasible."

Another remark was, "Tympanocentesis can be a very painful procedure and requires adequate anesthesia. If adequate anesthesia is not given, then it is unethical to perform."

So, there were some concerns. There is, I guess I would say, there is conceptually an agreement that it is

desirable to do tympanocentesis, but from a practical standpoint, is it necessarily doable, and to the issue of walking in the manufacturer's shoes, this is where we need to hear from people in the audience what you think is doable.

[Slide.]

So, what happened to the guidance? No change occurred to the guidance. That revised guidance has not reflected that change, and the clinical-only study is retained in the guidance at this point.

[Slide.]

Another point that came up as a question from someone in the audience, "Would placebo-controlled studies be valuable in establishing effectiveness for acute otitis media for regulatory approval?"

"If you select patients correctly, you will more likely have children whose acute otitis media is due to bacterial infections and therefore are more likely to see bacterial complications, and therefore a placebo-controlled trial would be unethical."

[Slide.]

Continuing along that, if a placebo-controlled study design was utilized, "you would have to exclude the very, very ill from the clinical-only trial, namely, those

with more pain, including a bulging and weeping tympanic membrane."

[Slide.]

So, to the guidance and the revised guidance, we have not changed. We are still recommending active-controlled trials.

[Slide.]

Subsequent to the advisory committee, we received a letter from Centers for Disease Control, specifically, the Drug Resistant Streptococcus Pneumoniae Therapeutic Working Group, that raised concerns to the agency about clinical trial design for this indication, and I have excerpted some of the highlights of text from that communication that was sent to the agency.

I do not believe that this communication came in light of the advisory committee, but I believe it came in light of some recent approvals where for otitis media there may not have been in vitro, as well as in clinical, necessarily the optimal coverage for the three main pathogens we typically see in this indication.

"Since clinical-only studies would need to be prohibitively large to detect a difference if one truly existed between two drugs and their respective pathogen-directed efficacies, the committee believes a

smaller bacteriologically-driven study would be more effective, especially since bacteriologic failure is correlated with clinical failure."

[Slide.]

The letter continues, to state that "We recommend repeat tympanocenteses three to five days after initiation of therapy (in 'a small number of' patients who were culture positive at baseline or only those deemed failures) as an important measure of treatment efficacy."

[Slide.]

What we have done in the guidance, and again we welcome your comments whether we should be going further with this, but we have recommended in the Analysis Section that the during therapy failure rates (study days 3 to 5) is recommended as a secondary efficacy analysis.

[Slide.]

We received one communication, one formal written communication from industry to this guidance document, and the next few slides capture those comments.

The first one is, "We do not agree with the requirement for tympanometry for the clinical diagnosis of otitis media. The procedure is technically difficult to perform in very young children, and the results are difficult to interpret accurately."

So, what we have done in the guidance is we have recommended that in the clinical-only study, that we primarily study children over 6 months of age. I guess I would parenthetically add to that, that in light of this issue of bulging TM, and where tympanometry is not necessarily needed for that, that one could see in the clinical-only study as the younger population, but there I think I would be mostly interested in documentation of the bulging TM.

[Slide.]

The next point from industry was, "The document suggests that 'recurrent otitis media' may be pursued as a separate indication from acute otitis media. Is this correct?"

The answer is yes, however, it is not dealt with in the guidance, and if this is something that a sponsor would like to pursue further, we recommend that you come and talk to us in advance, so we can work through the details on how the study should be designed.

[Slide.]

Another comment was, "We suggest that the late post-therapy visit for acute otitis media be deleted, as there is no need for clinical evaluations after the test-of-cure visit."

Initially, the guidance had recommended, I believe it was a four- to six-week visit. We agree. What we have recommended in the guidance document is to place this as an optional visit, and where there may be some value to the issue of middle ear effusion resolution if one wants to actually measure that, but it is an optional visit currently in the guidance document.

[Slide.]

In addition to the guidance, the following change is being proposed, and this is not in light of any discussion that occurred at the advisory committee, nor in light of any communication that came in from external stakeholders.

What we are proposing here is to include at study entry for children who have acute, less than 48 hours, tympanic membrane perforations and a swab of the exudate in the microbiologically-evaluable population as appropriate, but in general, it will be limited to the three main pathogens, Strep pneumoniae, Haemophilus, and Moraxella, although one could make an argument that it would be appropriate for Strep pyogenes, as well, in that Strep pyogenes frequently perforates.

[Slide.]

We have also added to the document a new section

which was not previously there, and it is an Analysis Plan Section. What we are recommending are two study populations primarily of interest, the per protocol. Again, there are two different studies, the clinical-only study and the clinical/microbiological study, and the intent-to-treat, all patients randomized who meet the inclusion/exclusion criteria as discussed this morning on the biometrics general discussion.

We would expect to see both analyses should show consistent results, logically consistent.

[Slide.]

From a primary efficacy perspective--and the next slide will be secondary efficacy--we were proposing two main primary efficacy points, which are clinical cure rate at the test-of-cure visit, and then also pathogen eradication rate at the test-of-cure visit in the clinical/micro study, but specifically in that analysis ignoring susceptibility to study drugs at baseline, being that for otitis media, it is essentially an empirically treated indication, although it is valuable to know what happens in those children where the susceptibility is either reduced or resistant, that to best mimic what is going on at clinical practice, that ignoring the baseline susceptibility, we believe is an appropriate way to go from the microbiologic assessment.

[Slide.]

Secondary efficacy. We recommend assessing clinical failure rates at the three to five day on-therapy visit, as recommended by CDC.

Pathogen eradication rate at test-of-cure where the baseline pathogen is susceptible to the study drug in the clinical/micro study.

Time to resolution of symptoms.

[Slide.]

Persistence of middle ear effusions at the test-of-cure or post-therapy visit if that visit, an optional visit, is obtained.

Clinical response by age group, children who are less than or equal to 2, and those who are greater than 2.

Clinical response for patients excluded from the intent-to-treat analysis.

[Slide.]

So, in summary, we are recognizing the desire for increased diagnostic specificity, and that is specifically for bacterial acute otitis media in the clinical-only trial. Active-controlled studies are still recommended.

The clinical-only study to this point, as we are proposing it, is retained in the document and the guidance.

[Slide.]

For the clinical-only study, that we primarily study children who are over six months of age with the caveat about, as I stated before, about a bulging TM, test-of-cure visit two to four weeks after study entry, and that the late post-therapy visit is optional.

[Slide.]

Everyone, I think we recognize, the advisory committee members from a year ago, consultants, CDC, everyone is concerned about Strep pneumoniae drug coverage for otitis media, and therefore, we want to put everybody on notice that in contrast to what had previously been communicated in the Points to Consider document, I actually saw this mentioned, this 25 in the pink sheet that came out on Monday, but the idea is that 25, this may not be enough for this indication where Strep pneumoniae is the major pathogen of resistance becoming more of a concern.

[Slide.]

Also, conducting studies in specific geographic regions is not specifically addressed relative to resistance. Time to resolution of middle ear effusions is recommended as a secondary efficacy analysis.

[Slide.]

To compare on-therapy failure rates as a secondary efficacy analysis, and to include acute perforated TMs from

a microbiologic standpoint within those first 48 hours, a swab of the exudates.

[Slide.]

That is the conclusion of my comments to you about what happened in the past, what we have done to the document, and I invite your comments and/or recommendations.

Thank you.

DR. CRAIG: Any specific questions on what he presented before we have the comments? Do you want to wait until discussion time?

Okay. Thank you, Brad.

We will move on, then, and Dr. Chesney will give her comments and then we will have discussion.

Committee Presentation

DR. CHESNEY: Thank you very much Brad. I didn't have tachycardia this morning, but I am having extreme tachycardia now, so bear with me. I often think I went into infectious diseases rather than general pediatrics so I would never have to deal with acute otitis media again, and it is one of those things that just keeps haunting you and haunting you.

What I want to review with you are some things that have happened since the advisory committee meeting where these changes were recommended, and I wasn't on the

committee at the time, although Brad has completely filled us all in on what happened.

[Slide.]

Obviously, acute otitis media is an incredible problem, it is an incredible expense in this country. Many of my adult colleagues blame us totally for the antibiotic resistance problem, and in response to this, the Centers for Disease Control convened the DRSP Working Group, I think it was in 1995, the Drug Resistant Strep Pneumoniae Working Group.

For the first couple of years the issue was looking as surveillance and what to do in the future, and based on the outcome of that committee, it was suggested that a TWG, Therapeutic Working Group, be formulated, and the first pneumococcal problem that the group decided to tackle was acute otitis media.

So, on March 20th and 21st of last year, many people from the CDC and many people that we would all recognize as being the leaders in terms of research in acute otitis media met to share ideas.

The specific issue was to review the therapy of acute otitis media in the era of increasing antimicrobial resistance, and the outcome, first of all, was the letter which Brad referred to, which was sent to Dr. Feigal from

the CDC, and the second written outcome will be a manuscript that has a number of recommendations that is currently in the review process.

What I wanted to share with you today were just some of the recommendations and suggestions from that committee. I spoke with Scott Dowell yesterday who is the main author of the manuscript, and I am comfortable sharing these with you.

Historically, antibiotics have been selected for acute otitis media based on their beta-lactamase activity, their palatability, their cost, and convenience of dosing, but we are now really in a totally different era.

Dr. Edwards from Nashville, who is doing an otitis media study in a number of offices there now, came to speak to us a week ago, and unlike the 40 percent resistance that we are all seeing with invasive isolates, she is seeing a much higher resistance in isolates from otitis media.

So, we now need to really focus on the in-vitro activity against drug resistant Strep pneumoniae and the in-vivo ability to eradicate these organisms. On top of this we need to consider when we add new antibiotics that we are not making the problem of resistance worse.

[Slide.]

The letter from the CDC to Dr. Feigal pointed out

that we need to focus now on the pneumococcus in part because it is the least likely organism to resolve spontaneously without antimicrobial therapy. Children with non-typeable Haemophilus and Moraxella otitis media will frequently resolve on their own, and you probably know that in several Scandinavian countries now, they don't treat acute otitis media with antimicrobials, and it is because the Moraxella and H. flu do resolve pretty much on their own, but pneumococci don't.

As the treatment of acute otitis media is almost always empirical, any agent approved should have documented microbiological activity against all three major pathogens, and Brad mentioned this.

There was great concern on the committee that there have been two antibiotics approved recently that did not have in-vitro activity against pneumococci, and it is in very small print in the advertisements and journals that most people wouldn't pick up, most clinicians are using these two drugs for all three pathogens, not recognizing, so the committee felt strongly that any subsequent drug approved for acute otitis should have activity against all three pathogens, and that clinical-only studies show similar effectiveness between agents despite substantial differences apparent if you do tympanocenteses, and the reason for this

is--and this is what we call the "biocreep" phenomenon--80 percent of acute otitis improves on day three to five without any treatment. Unless you use strict entry criteria, patients with middle ear effusions that are uninfected are included, and the criteria for cure improvement in the past have been based almost exclusively on clinical outcome, and not tympanic membrane findings.

So, clinical-only studies may be missing what is really going on.

[Slide.]

To pick up subtle differences using clinical-only studies, you would need huge sample sizes, whereas, you could use a very small number of patients if you were looking at microbiologic outcome where the important differences in efficacy would be more easily detected.

Again, the committee felt very strongly that evaluation with repeat tympanocentesis at day 3 to 5 is an important measure of treatment efficacy, and that repeat tympanocenteses, assuming an initial one was done, could be limited to treatment failures, and this is not from the committee, but my thought that patients with non-susceptible pneumococci on the initial tap might be good candidates for a repeat tympanocentesis at day 3 to 5.

This is a quote. "Most participants suggested

that a carefully conducted repeat tympanocentesis study be considered a critical element for approving antimicrobial agents for an acute otitis media indication."

[Slide.]

There was one important dissenting viewpoint, because this is an individual who has had a great deal of experience with otitis media, and this individual felt that large clinical trials could be continued if stringent entry criteria were used, stratification was done by clinical severity, and outcome measures included a symptomatic response, but also otoscopy findings that tympanocenteses were done for failures and that there was in-vitro evidence of adequate coverage for all pathogens, as we mentioned already.

The suggestions. All agents approved for acute otitis media have acceptable activity against all major pathogens--which we have already talked about--and would all major pathogens now include resistant pneumococci, not just pneumococci.

This, as I mentioned, was felt to be documented most easily with trials using repeat tympanocentesis after three to five days.

[Slide.]

And that the trials could be restricted to small

numbers of children for whom strict entry criteria were used, the pathogen was identified on an initial tympanocentesis, a follow-up was done again, maybe based on treatment failure or if the initial tap showed a resistant organism, and stratification by clinical severity.

[Slide.]

In summary, future studies, it might be considered that a small number of children be entered into tympanocentesis studies to document bacteriologic eradication.

Many people on the committee felt that the double-tap study was preferred. By that, they mean an initial tympanocentesis and a follow-up without restrictions of it being a failure or an initially resistant organism. But they would be willing to limit that follow-up tap to those with a failure or--again, my addition--the non-susceptible organism initially.

This is not from the committee, but it would be an ideal time to obtain drug levels from the middle ear fluid at that repeat tympanocentesis, and I think any future study, it should be considered that a minimum number of patients with a non-susceptible organism be considered.

I know with the IM ceftriaxone study, there were a very small number of non-susceptible pneumococci there, and

so the indication was only for susceptible.

The clinical-only studies, I think have to be continued in terms of tolerance and safety issues, but again a concern is that an appropriate comparator drug be selected, and these are the three drugs that I think most people now recognize are good for the resistant pneumococci.

[Slide.]

The last overhead. Obviously, shorter courses should be considered, and this again is an aside, but in this era where we are promulgating judicious antibiotic use, that it be considered that the normal flora be looked at before treatment and at the end of treatment at some period of time to see what the issue is in terms of inducing antibiotic resistance with this particular antibiotic.

Thank you for your patience, and I hope I have represented the committee's deliberations well. I think I will stop there and see if people want to comment.

Committee Discussion

DR. CRAIG: I guess we can take your specific comments and address them one at a time. You didn't have any specific questions, Brad?

DR. LEISSA: No. I think our overall question to the advisory committee members is what, if anything, is wrong about the current guidance in terms of what should be

changed, because this is kind of like, you know, we are getting closer and closer to a "final guidance," and if there are things that should be changed, we need to hear about those now.

DR. CRAIG: Carl.

DR. NORDEN: I guess I will start and try and be slightly provocative. Based on what Joan just presented, which I think was very clear and represents input of a fairly expert group, why would we want to do clinical-only studies any further at all for this indication, because I think there are problems that will occur with them, and I think again if you take a disease that has an 80 percent or whatever the percent is spontaneous remission rate, I don't see what information we are going to gain.

DR. LEISSA: Prior to 1992, when the whole issue of a clinical-only study was first entertained by the Division, within about a year or two prior to that, we had a number of sponsors that came in to us and said we understand that to date, we have always wanted tympanocenteses, microbiologically, clinically done studies for otitis media, and I am paraphrasing what they said to us, but the idea was that we are concerned that if you want these studies done, we are going to have to go overseas, we are going to have to go to Central America, wherever, where it is easier to get

patients because we are not finding investigators who can do tympanocenteses or are willing to do tympanocenteses, so I think at that point there may have been an overriding concern about having the patient populations, the children not necessarily being of North America, that is historically.

So, in light of the comments that came from the CDC and the Working Group, all I would go back to, I think the issue that I am mostly interested in is a practical one, which is to say what do companies think they can do, because again, if they don't say anything here, we typically hear about it later from them, which is when they come in and they say to us we just can't do these studies here in this country.

So, I think that is an overriding issue that I need to have some sense of, is how doable is this for the sponsors who are doing these trials. If we say tympanocenteses are going to be the recommendation for all future drugs for otitis media, and also the point that we are talking about, Dr. Chesney and I earlier, which is that where some drugs that are first developed, they are developed initially mostly for adults, and then when they go into pediatric populations, sometimes that is limited to the otitis media population, and if you do a smaller

clinical/micro, where you may be losing something on the safety experience with the product in that, so that is another side to that.

DR. CRAIG: Go ahead.

DR. BLACKWELDER: It seems to me that for a comparative efficacy, such a study is of doubtful value because, first of all, there are a large number of spontaneous resolutions, and second, if you can't be confident that it is highly specific for bacterial infection, those are at least two, maybe others, which would tend to make you find drugs similar even if there are some important differences.

DR. LEISSA: Specifically, your comment is to the clinical-only study, is that correct?

DR. BLACKWELDER: Exactly.

DR. CRAIG: The question I guess in my mind is I think what you have tried to do with some of the changes with looking at otoscopy and things like that is you are trying to tighten up the diagnosis that you are dealing with, bacterial otitis media, but what I still see as a problem is to try and identify those patients that will have a very high spontaneous cure where the antibiotic is not very helpful.

There have been a variety of at least one major

placebo-controlled trial recently that showed reasonably good results compared with antibiotic therapy, but again what they were doing is those I think were kids that were over the age of two.

The question I have and the concern I have with many of the trials is maybe it's that population from six months to two years old that is the population where we are going to have the greatest chance of seeing clinical differences because it is in that group that probably the antibiotic is playing a role.

Yet, if you look at the studies that have been done in the last few years, and look at the mean age, the mean age is about four and a half, so that we have been looking at an older age group population where we are essentially diluting out that population where we might see a clinical difference.

My thought would be that maybe we need to have a certain percentage or a certain in the clinical-only trials of patients that are in that lower age group, where the antibiotic may be important and one might be able then by using clinical means alone, be able to show differences. But including all the older ones where it is very unlikely for that to occur, I think doesn't let us make some of the same mistakes that people feel have been made already in the

approval of certain ages.

Yes, Dr. Henry.

DR. HENRY: I agree with what Dr. Craig said about looking at the younger age population. It seems like now in infectious diseases, you know, we certainly go around giving out advice about when to use a drug, but it seems like we are also giving out advice about when not to use a drug because of the abuse of antibiotics especially in the pediatric population where parents want antibiotics in hand when they go on vacation in the summer.

So, I think we really do need to know something about the microbiology because there may be times when we don't need drugs, but when we need them, we need to know that they are effective against the drug uses in Strep pneumo.

Now, when you get into the younger age population, there is also the dilemma of, you know, if you want to prove that there is an organism there and know what it is, and if you do tympanocenteses, what is this argument about having to have anesthesia, so that it's an ethically correct procedure, and what studies are going to be done.

You know, if you have to have someone there who is going to anesthetize the kids, so you can get an appropriate specimen, that brings in another element that makes it much

more complicated, which goes back to Dr. Craig's concern or question and comments about maybe in that younger population you could do a clinical-only because these were kids that, you know, how many of these kids are you going to anesthetize. I mean it's very complex, and yet at the same time, I want to know if I am prescribing a drug that is going to work against drug resistant Strep pneumo and in the right age population, but do I want to do a tympanocentesis without anesthesia? I don't know. I mean that is something that is going to have to really require a lot of thought on the part of industry, as well as the FDA and us.

DR. CRAIG: I think to me, the age isn't as important if one is doing repeat punctures and finding out what the organism is, but when you are not doing that, and you are doing the clinical-only study, then, I think you want to try and look at it in that population where you stand to see a difference and the antibiotic is going to be most beneficial instead of looking at it in the group where it adds such a little bit that it would be hard, you just need such a huge number of patients to try and sort that out, but where punctures are, I mean then I think you can, at least in my mind, you are looking there at the bacteriologic response.

You have added in a three- to five-day clinical

response as being one of the secondary endpoints. I know that Ron Dagan has done a study to show that if you look closely at kids at that time period, you can see a difference, but that was with a whole scoring system and everything that he used.

I am not sure with the kind of information that you are gleaning that you are going to be able to actually pull anything out. I mean it is a possibility, but I am not sure that with the kind of information that is currently obtained, without scoring it in some way, that you would be able to pull out some differences, and that would be my one concern about having that be a clinical determination at Day 3 to 5.

DR. LEISSA: And the idea with the three to five is there is greater value when it would come to repeat tympanocentesis, but to the issue of using it as a valuable clinical endpoint, that is where you are lesser.

DR. CRAIG: A question for the FDA. How often--I know right now you have some recommendations, I see in clinical trials where failures are requested to get re-taps of tympanocentesis--what percentage of those actually do you get a tap, 50 percent of the failures?

DR. LEISSA: The percent of the children that come into the study, what percentage of those are failures?

DR. CRAIG: Of those that are failures, that are clinical failures, get re-tapped?

DR. LEISSA: Very small numbers. It is hard to determine at what point that failed, whether the issue was that the investigator just didn't feel that it was of value to do, what was the incentive for the investigator to do it, whether the parent refused to have it done, but it is very, very infrequent that we get that typically done.

DR. CRAIG: Even though it's--

DR. LEISSA: --even written into the protocol.

DR. CRAIG: So, that is one of my concerns.

Yes, Dr. Reller.

DR. RELLER: In listening to Dr. Chesney's wonderfully succinct presentation, I wonder if we are deluding ourselves about objective assessment of safety and efficacy.

If drugs for otitis media are approved with clinical trials alone, with an agent that is not effective against Streptococcus pneumoniae, and even that is not demonstrated to be efficacious against Strep pneumoniae that are resistant to penicillin, when one thinks about the implications of approval for 25 million prescriptions a year in this country, and all of the concerns on both sides of drugs that are--the fear of resistant pneumococci in using

drugs that may or may not be necessary, and then those that are assumed to be efficacious because they are approved for otitis media without reading the fine print, and they have never even been shown to be efficacious against perhaps the principal organism that needs to be treated when the disease is truly present.

So, I would suggest that approval of a drug with clinical trials alone should not be possible, and maybe the 25 should not be a debate about 25 Streptococcus pneumoniae, but rather--unless there are 25 strains documented to be present and eradicated that have an MIC above 0.06 as a minimum criteria for approval for a drug that is purported to be efficacious in the therapy of acute otitis media.

DR. LEISSA: Dr. Reller, when you are saying the MIC of 0.06, is that penicillin?

DR. RELLER: What I am talking about are the intermediate resistant. The strains that we are concerned about are those that have penicillin MICs above 0.06 for Streptococcus pneumoniae, and without an organism, without tympanocentesis, we don't know whether the drugs that are being developed, whether they work, and whether some of the ones that are promoted, whether they are necessary.

DR. CRAIG: Dr. Chesney.

DR. CHESNEY: Two points. One has to do--I am

going to remember one, I hope I remember the other, so-called senior moment--Bill makes a superb point about the age that should be studied, because the resistant pneumococcal problem is primarily in children under two, and particularly in children under six months, and Dr. Schwartz sent some comments back to Brad and he said why would we cut off at six months, when the children under six months are equally as at risk for these resistant organisms, and the second point I now recall, which is probably the most important and the most emotional of all, is why don't we do tympanocentesis, because we worry about the pain, we worry about the skill involved. I am terrified of sucking the malleus up into a tympanocentesis needle when I do one, and it is so easy to do the clinical-only. I mean, you know, without really strict entry criteria, you could put every child who has got a red ear and fever on the drug, and get it approved--no, I don't mean that--I mean get the data.

But doing a tympanocentesis is a big deal, and I don't think any of us means to minimize it, and many of the people at this meeting are those people that do them routinely, and I think that if we all learned how to do them and if we learned the techniques for rapid and quick anesthesia, that there would be much more of a comfort level with it, but I think that the benefits to be gained for the

child are enormous. To take a child with a red, bulging ear who may have an organism in there with an MIC of 16, and put them on trimethoprim sulfa to which it is totally resistant is not doing the child a favor, and that is more in terms of treatment, and not study, but the point is the same.

DR. CRAIG: Other comments?

DR. HENRY: I just have one comment. When the Therapeutic Working Group was coming together, or maybe those were just your comments on that last overhead about looking to see if the normal flora had changed in its susceptibility, I mean were you thinking of doing nasopharyngeal swabs or what normal flora specifically were you including in that?

DR. CHESNEY: Well, that was more my comment. Although we discussed it at the committee, it is not an official recommendation. I have seen a study that came out within the last few months that did look at pneumococci in the normal flora at the beginning of treatment and the end of treatment.

DR. HENRY: But by getting nasopharyngeal swabs?

DR. CHESNEY: Yes, a nasopharyngeal swab. So, that is just a suggestion.

DR. CRAIG: Is there a way that you can tighten up, so that you do get punctures on the failures?

DR. LEISSA: Again, I don't understand the phenomenon well enough to why we are doing so poorly when it come to the repeats in tympanocentesis.

DR. CRAIG: That clearly could be a population which would be useful, and furthermore, I mean theoretically, if you have got another drug in your pocket, those would be patients that if they fail, and you have got a resistant organism, they might be perfect patients for the next drug.

I mean I think that was the one of the feelings that I got from the CDC meeting, was that doing taps on failures would be one way of doing it, and there are investigators out there specifically that have generated a large number of patients doing that specific procedure.

So, I think that is a reasonable alternative at least initially. It is sort of one of these things that as you gain more information, you get more results. You can then start making better decisions, but I agree that I think it would be difficult outside of a few people to actually get routine tympanocentesis.

So, my feeling, if that is difficult to get done, what you try and do then is tighten up your inclusion criteria to try and make your clinical-only study to really involve those patients where the drug probably has got its

best chance to show benefit.

So, I would try in some way to ensure that you have a significant number of children under the age of two in those studies, so that you have a better chance of having--if there is a difference to be seen--of seeing that difference between the new drug and the comparator.

Otherwise, if we keep using older children, I don't think in the clinical-only study you will be able to see differences.

DR. LEISSA: When the Points to Consider came out in '92, the way people were I think interpreting that was that there was going to be "the one large clinical study," and the "smaller, micro study," but I think also if we look at this issue about the number of Strep pneumoniae that we are actually interested in getting, you know, true information about in terms of tympanocentesis-driven for Strep pneumoniae, and to the issue that Dr. Reller mentioned about looking for actually when you would come up to a number of actual documentation in the non-susceptible Strep pneumoniae, the microbiology study is no longer "a small study," which I think we are all looking for.

So, maybe we should have some confidence that putting a different standard to the issue of the microbiology study, will actually get us a lot more

information than we may be getting currently based on what the Points to Consider was implying earlier.

DR. CRAIG: Can there be also some incentives to actually do some of the re-taps, for example, if you do re-taps, you can get by with a smaller number of organisms in order to get approval if you can demonstrate clearly that you had the organism at the beginning and at the end or at four to six days, which is usually the time period, the organism is gone by showing eradication that would allow you to have a smaller number of organisms to get approval against resistant organisms as compared to if you only had ones in which you had a tap at the beginning and only clinical results after?

DR. LEISSA: I think the incentive for doing the re-taps hopefully would be where you were able to actually have in the face of what appeared to be clinically lack of response, but actually had eradication, and then with subsequent follow-up after that, that there was no recrudescence recurrence of the organism.

There you have a document of persistence, which otherwise in the position where you don't know, you would take the conservative approach in saying that that would be a presumptive persistence, and that may go towards the numbers that might support a labeling of an organism

especially when it would come to the issue of decreased susceptibility.

DR. CRAIG: The problem you have with otitis as far as getting a final, I mean the longer you wait, the less chance there is for fluid to be there to re-tap, and that is why I think many of the people have used around four to six days, because if you waited out all the way to 10 days, the number that you are going to have something to re-tap clearly drops off.

Yes. Comments from the audience? Yes.

DR. HAFKIN: Barry Hafkin from Pharmacia Upjohn.

I am in the business of trying to find investigators to do these studies, and I can assure you that trying to find investigators that will do one tap is possible. We might be able at any one time to find 10, 11, maybe 12 centers, but the number of American sites that will do two taps may be counted on one finger. At least that has been in my experience, and that one site is not very productive.

I would want the advisory committee to realize that it really is very hard to get conscientious pediatricians, even those committed infectious disease, to do that first tap. The second tap, we literally have to go abroad. There are sites abroad that will do it.

So, if the advisory committee would want that kind of data to be used for registration, then, that is possible, but I would want to warn you that we would be using primarily extra-territorial data.

DR. CRAIG: How about improving the second tap in failures, at least getting that?

DR. HAFKIN: Again, it is not a matter of my not being willing to pay for it, it really isn't. I mean there is no benefit to me not to provide that data to you, because we are all interested in that.

I mean at the end of the day, if you help me register a drug that doesn't work, it doesn't do very well for my company in the long run. It certainly doesn't help the patients. We would love to get that data.

DR. CRAIG: Do you think in an era with more resistance now that it might be easier to do as compared to the past?

DR. HAFKIN: Well, I think that the number of people who are at least speaking to the issue of maybe re-tapping kids that haven't done well is possible, but still you have to remember the interface between the parent, the doctor, and the patient is real.

I mean this is an unhappy child, there has already been one tap, you know, the mother has been petrified, the

clinician doesn't want to hurt the child, it is just not that easy, and the reality is that we are tapping people who again may have no viable bacteria, but that persistent effusion, it may have been a bacterial infection, but it may have nothing to do with the residual bacterial infection.

Let me make a couple other points, just brief ones. It amazes me. Now, I am an adult ID doc, so maybe I can't empathize with one pediatric aspect, it amazes me that we would want to prove that antibiotics work against bugs that don't cause disease, like Moraxella and Haemophilus, it's not typeable, and every time I hear us talk about that as a community, as an infectious disease community, I truly don't understand it.

I mean much of the world does not treat that syndrome of otitis associated with Moraxella and Haemophilus that is not typeable, and why we would continue--you know, this has been coming up for years, we would have got to have antibiotic therapy to cover bugs that probably shouldn't be treated anyway. It is just a wonder to me as an adult ID doc why would we want to do that.

The third point I would say is that I think that there is value in the clinical trial. Remember we are not only talking to you at the end of the day about a study that shows efficacy, but safety is an important issue. Let's not

throw that away. And being able to say that you can give that antibiotic to hundreds of children and not have diarrhea, not have rash, not have fever is very worthwhile.

So, what I would urge you to do is to remember that safety is important, it is very helpful to us as an institution, you know, in the business of selling and making drugs. We think it brings value, and that idea that balance between a small microbiologically driven study and a clinical study, I think makes good clinical sense.

Thank you.

DR. CRAIG: Thank you.

Go ahead. Next.

DR. HOLLY: Hi. Preston Holly from Glaxo Wellcome.

I would like to reiterate many of the comments that were just made, primarily the ones about the difficulty in finding investigators to do tympanocentesis in the first place, but to do a double tap is again almost impossible in the United States, and even outside the United States it is difficult to find such investigators.

One of the reasons investigators have, in fact, in studies that we have conducted in acute otitis media with effusion where part of the protocol was to tap patients who failed on therapy, we got very few patients that actually

had the second tap. We did have a few, but the primary reasons the investigators give us are that after the first tap, three days later there is really not much fluid even if the patient looks like they are failing clinically, there is just not enough to tap.

Remember we are talking about children with very small tympanic membranes to start with, and the concern over the child's welfare is one big concern. Secondly, in patients who are improving, I don't know of any investigators in the United States, but there might be a few, that would be willing to tap those children who are improving to show that the organism is gone, and to get that through an IRB might be very difficult also. It really is unethical in my opinion.

So, we are left with very few sites that can even do these studies, and then if you put on top of that the requirement for more than 25 Strep pneumoniae, does that include, as Dr. Reller suggested, 25 isolates that are resistant to penicillin? When you get down to actually looking at the numbers, of the numbers of children you enter into the study, the numbers of children that then have Strep pneumo that is cultured out, and then the number of children where that Strep pneumo is resistant to penicillin, and then the number of children who fail and come back, and the

parents are willing to have them re-tapped, and the investigator is willing to do that, it would take huge studies to do this type of work, to get the numbers that I think are being talked about here.

So, again, I would say that by making these studies more restrictive, it is going to be more and more difficult to get drugs through this process or even have manufacturers consider putting a drug into the process.

On the other hand, I think the points that you have made, and others, about maybe stricter criteria for the non-tap studies are good points, and those could certainly be addressed as we would agree that might be a good way to go.

Thank you.

DR. CRAIG: Thank you.

DR. HOPKINS: Preston, there is one site in the United States that will do second taps, but I am not going to tell you where it is.

[Laughter.]

DR. HOPKINS: Scott Hopkins from Pfizer.

Just to put some numbers to the thought that Preston Holly just expressed, we did a trial with azithromycin a few years ago that involved single taps at baseline, and we enrolled a little bit over 300 patients

into that trial. About half of them had positive taps, so we have 150 on our drug and 150 on the comparative agent.

Half of those 150 on azithromycin were positive, so that is 75, and that broke down about equally into the three predominant organisms. So, that means that we had about 25 Strep pneumos and just barely met the criterion at the time.

I think at the time the prevailing incidence of intermediate and high level resistance was on the order of 10 or 15 percent, so one would have expected that we would have gotten two or three of those in our 25 Strep pneumos. We were unlucky, in fact, as any statistician could tell you that we might be, and we didn't get a single one.

So, we could have even doubled that study and tried to enroll 600 patients and get taps on all of them, and we still might only have ended up with two or three or four subjects who had organisms that were really of interest to us, the resistant Strep pneumos.

I don't think any of us really believe that we can make good judgments on the basis of two or three or four organisms in terms of whether or not a drug is really working. This is the problem that we have with pen-resistant Strep pneumo and low-incidence organisms in any sort of clinical trial situation whether it is

meningitis or community-acquired pneumonia or otitis media, is how to deal with these, and I think, unfortunately, we are in the situation, or perhaps fortunately, of having to go back to the other sources of information that we have, that is, the in vitro and the animal models, and so forth, and putting more reliance on those when we don't have the clinical information that is available to us.

DR. CRAIG: Thank you, Scott.

DR. WICKLER: Matt Wickler from Bristol Myers Squibb.

I won't beat tympanocentesis to death, so I will bring up some other issues, but I do agree with everything that has been said, and I know 18 people who do tympanocentesis and do good quality work that would be acceptable to the FDA, and I know one who would do repeats, and that's it.

When you have all these companies trying to vie for the same sites, you can imagine what a contest it is to try to be the first one to sign up your site, so you get your study done for the coming respiratory season.

I also want to second--this came during the resistance meeting yesterday, and that is trying to use animal models, other things that we feel are predictive to help give us an idea of what will and won't work, because I

think trying to depend upon tympanocentesis in clinical studies, that really would give you the answer, it may be not be a practicality.

I want to discuss two other things. As far as the proof comparators that were mentioned, that came from the CDC Working Group paper, a draft I saw, two of the three drugs that were mentioned were at doses that are not approved by FDA labels. They are actually higher doses or more frequent doses. I don't know how the agency would deal with that.

The third issue is you mentioned looking at time to resolution of signs and symptoms, I believe, and the question is how do you do that, do you call up the patient every day, do you make them come to the office every day? I think it is a good endpoint and it is valuable.

I think we have to give some consideration as to how you actually do that, and mechanisms on how you actually do a study to actually get that sort of information.

Thanks.

DR. CRAIG: Thank you.

DR. YEADON: I am Arnold Yeadon. I am a self-employed consultant. As you can probably tell, I don't come originally from the United States.

Just a couple of points. As I heard the

discussion, I was a little concerned that people have said, well, if you can't do tympanocentesis in this country, you can do it abroad, and it seems to me that if is unacceptable or unethical to do it on American kids, it is equally unacceptable to do it on foreign kids, even Brits.

[Laughter.]

DR. YEADON: But the other practical concern that I have is maybe the tympanocentesis is itself potentially curative. I am 67 years old, and when I was a kid, I had serious otitis media, and it was treated by what we called in those days myringotomy.

I remember it hurt like hell, and a lot of pus came out, and then my mother treated it with hydrogen peroxide, but I got well anyway, so maybe by sticking needles in people's ears, you are somehow interfering with the question of whether your antibiotic works or not, and maybe whether it is even needed or not.

By the way, I wear two hearing aids now. Thank you.

DR. CRAIG: And there are placebo studies looking at that, and obviously, there is some natural eradication that occurs much less with pneumococci, but fairly significant for Haemophilus and Moraxella.

Any comments? Yes.

DR. GOLDBERGER: First, to address that comment about using the doses higher or more frequent. I think from our perspective, the issues would be I think relatively straightforward. One is ensuring that there was sufficient safety information that the dosing regimen was okay, and I would expect in most cases that would be the case.

The only other issue I can imagine that would be a problem is if the more frequent or the higher dosing led to increased discontinuations due, say, to GI or other toxicity. That might prove a problem.

But if you are attempting in general to compare an experimental drug to a standard drug, and you using the standard drug at a higher dose, presuming it is well tolerated, that ought not to pose too many problems except to the experimental arm.

My only other observation is before we close this session, now that the committee has had the opportunity to hear what the industry has had to say about some of these trial issues, if we could see if there are any more comments about perhaps the guidance document or any changes, et cetera.

DR. CRAIG: First, to answer that question, I guess I would respond still that I think that I would try and tighten up the criteria for the clinical-only study and

possibly, I mean even before the tap ones. I mean I would try and make sure that there was a significant percentage of the patients within a certain age.

The reason I guess I wouldn't say entirely less than two is because what the FDA frequently does, then, is it labels it only for kids less than two. So, I think it would be good to have a few of the older kids in there, but I would want to make sure I had a significant number of them at the lower age where I would have a chance to see clinically if the drug would work, because I think if we have got this problem of getting that kind of data, then, I think we have to tighten up the study population to identify a population where we might see a difference.

Secondly, I would try and at least in certain situations, see if it is possible to obtain some incentives to try and get some more data, such as in failures, and also for the one investigator that does do the re-tap studies in the United States, so that it may be that you can get an approval earlier.

I mean this may be a career development area for, in fact, pediatric infectious disease people in the future to learn how to do double taps, so that we can increase the number that are able to do it.

But I would try and at least in the criteria, try

and make some incentives to get the better data, so that there is an incentive to try and obtain it.

Other comments? Anybody else? Joan.

DR. CHESNEY: I would just make a brief anecdote here. When I was a young faculty member at the University of Wisconsin, and Dr. Craig and Dr. Cal Kunin were two of my mentors, I was at a conference one day and Dr. Kunin challenged me because I had not put a needle in the bone of a child with osteomyelitis to attempt to retrieve the organism, and he said, "Why didn't you do that?" And I said, "Well, I didn't want to hurt the child," and ever after that, he made fun of me, not wanting to hurt a child, but that is intrinsic to pediatricians, and I think that we really have to learn.

I think a career development issue is that we need this information. I think that is the message from the committee, is that we badly need this information, and I think we, as pediatricians, need to figure out how to do the anesthesia required and learn the procedure, so that we can provide more sites.

DR. CRAIG: Dr. Henry.

DR. HENRY: I just have one question that came up in Brad's presentation that I guess I would like Barth to address, and that is using an ear swab from a child who has

a perforated TM and doing a swab on what is sitting in the ear canal.

Would that be acceptable as a microbiologic specimen from the standpoint of a microbiologist?

DR. CRAIG: A relatively recent, it was within 48 hours, acute?

DR. LEISSA: Yes, an acute perforation, right.

DR. HENRY: Acute perforation, an ear swab with an acute perforation.

DR. CRAIG: Limiting it to the three organisms that we are talking about, in other words, you don't normally have those organisms as part of the normal flora of the ear canal, do you?

DR. HENRY: You shouldn't, but I am just wondering if there is going to be so much overgrowth that you can't interpret even what is there. I mean it is true, whatever is sitting in the canal is going to be contaminated by what was there before, and maybe that is not pathogenic, but I am not certain how reliable that would be, and you brought that up at the end.

DR. LEISSA: Right. The issue would be practically is if the child came into the study, and they had the perforated TM, and they had clearly a purulent exudate coming from the ear, and that swab was sent out for

culture, and after going through isolation, you found Strep pneumoniae, whether you would say, I just don't know if that was truly causative or not, and whether we should accept those as being part of the definitively microbiologically evaluable population.

DR. CRAIG: I would accept it.

DR. LEISSA: Would you accept it for the three, as well as Strep pyogenes?

DR. CRAIG: Yes.

DR. LEISSA: Haemophilus and Moraxella?

DR. CRAIG: Dr. Reller.

DR. MURRAY: I think it gets back to one of the other questions about should we be accepting Moraxella to begin with even from the tap.

DR. LEISSA: Dr. Reller is rubbing his head.

DR. RELLER: I would say I have just been converted to clinical-only studies.

Seriously, if one looks at, for example, nasopharyngeal specimens in blood isolates, the colonizing organisms are at least as resistant, if not more so, than the invasive ones when this has been looked at, but I think that if one had an acute perforation like happened in Britain 60-some years ago, and you recovered a Streptococcus pneumoniae or a Strep pyogenes, I think they shouldn't be in

the external ear canal just after coincidentally a clinical entity that perforated.

The Moraxella and the non-typeable Haemophilus, and all that other rubbish, I mean I discount those. It is a Group A Streptococcus and a Streptococcus pneumoniae. I think that would be a reasonable thing.

I would like to just have the opportunity at the conclusion, when one looks at this whole discussion, you know, some numbers are missing on the order of 20, 25 million people treated, is this the correct number, Dr. Chesney, a year, and rates of resistance widespread in every community in the United States of 15, 20, 40, 50 percent and growing, with a disease that may be helped by relieving the pressure, we may end up with a tightened clinical definition, but I don't think at the end of the day we have any evidence for efficacy of what people are most concerned about, and something just doesn't add up to me as a criterion for critical trial design, trying to answer the questions on which we know once a drug is approved, is used more widely with less stringent criteria, without regard to the issues of engendering resistant organisms with unnecessary treatment, et cetera.

I mean I think that we have an opportunity to do good with science, with smaller numbers, that can escape us

if we just say, you know, they got better and it's okay, and it's as good as everything else that is currently overused.

Dr. Chesney, what I haven't heard, and I am not a pediatrician, I have never done a tympanocentesis, but with the training that could be resurrected perhaps, what are the recognized, is this truly a dangerous procedure?

There are a lot of things that are done in medicine that they certainly have some risks, but they need to be done for ultimately the benefit of often the individual patient--you know what I am trying to say--how dangerous is this? I mean the implication is that this is so dangerous that nobody will undertake it. What are the facts?

DR. CHESNEY: My understanding is it is not a dangerous procedure at all. The problem is immobilizing the child and relieving the pain, and we were just talking about this issue, there are ways of relieving the pain, there are ways of giving a short-acting, not analgesic, anesthetic agent.

The thing that the children like the least is being restrained, and they have to be restrained. You have to have that ear steady, so you can put the needle in, but if the child is restrained, and if you have used adequate pain control, my understanding is that it is not the least

bit dangerous. It is known where you go in and ENT surgeons do them every day, and many pediatricians actually do them every day. It is a selected number who are very comfortable with it.

DR. RELLER: Not to belabor the point, but down the line it seems to me that the distinction carefully between what is dangerous and what is difficult is very important, because there is also a balance between what is valuable and what is difficult and what is worth investing the effort in order to get an answer that is extendable for general usage.

DR. CHESNEY: I think one other point is that often the pain is relieved when you do the tympanocentesis, as was mentioned. Once you put the opening in the ear canal, and the pus comes out, then often there is a great relief of the pain and the pressure that was there. Just an aside.

DR. CRAIG: Two last comments from the audience, I think, and then we are going to need to leave this subject.

DR. LEROY: Bruno Leroy, HMR.

I have a question regarding the number of pathogens obtained outside the U.S. territory. It seems the question is to know the presumed eradication of a pathogen of a certain MIC. Is there any scientific rationale not to

accept foreign data?

DR. LEISSA: There is no scientific rationale other than to say can it be extrapolated to the U.S. population, so that depends, but we certainly do accept foreign data all the time.

DR. LEROY: It seems the question is not the comparison, but here the eradication of this pathogen of a given MIC. This MIC, a foreign MIC of the same value in the U.S. territory, you can extrapolate the data. So, the 25 number you obtain, the number of 25 isolates, for example, you obtain.

The extrapolation will depend just on the value of the MIC. If you extrapolate in South Africa, an MIC of 2 with the compound, with the same method of treatment, it will be valid whatever the location.

DR. CRAIG: To me, the resistance mechanism of the organisms in Europe is the same here. I mean where I could see it could be difficult to extrapolate might be for macrolide resistance from Europe, which is primarily MLS, while in the United States, it tends to be much more an efflux mechanism. So, obviously, what works against an MLS may not work against an efflux, so one would have to have that data, as well, but in terms of penicillin resistance, I would sure think it should be extrapable.

DR. LEISSA: I think the only issue we have ever had, and this applied, for example, back to Haemophilus influenza a number of years ago with the issue of beta lactamase-producing organisms, and some countries are more effective culturally than we are in terms of limiting antibiotic usage, and therefore they tend to have lower prevalence of these organisms, and whether in your database to represent the drug's activity, and the indication that you would have a lot of patients in areas where resistance wasn't as much of an issue, and whether in those clinical-only studies you would be using that extrapolate to in this country, where we might have high resistance. I think that would have been and is the concern.

DR. ALBRECHT: Let me add to that by briefly mentioning something I will talk about tomorrow, about foreign studies. Yes, the Code of Federal Regulations does have criteria where foreign data are acceptable for registration of drugs in the U.S. marketplace, and it deals with--I mean you bring up the issue of MICs, and clearly we need to have the other elements, as Brad alluded to, as far as applicability, so we would look to make sure that the children, the character of the children, the underlying diseases, if any, the other sort of socioeconomic aspects are applicable to the U.S. population, that the

microorganisms, as you mentioned, the MICs, are applicable, that the study was conducted in such a fashion that it is analogous to what we would see here, so that we could make the assessment.

Lastly, of course, we do request that there be access to the patient data. So, if those kind of elements can be met, then, we have actually used foreign data in granting approval for agents for--I can't right now recall if otitis, but certainly in other indications.

DR. ALEXANDER: John Alexander from the FDA. I actually had a question for Dr. Chesney. In your presentation you had mentioned that the CDC Working Group had said that they would be interested in seeing data on drugs that were only effective against all major pathogens, but everybody here seems to be most concerned about Strep pneumo, and not concerned as much about Haemophilus influenzae and Moraxella, so if there are drugs that were being developed that were specifically active against the gram-positives, and not necessarily as active against Haemophilus influenzae and Moraxella, how would you treat that drug?

DR. CHESNEY: Well, I can't speak for the committee, but I think they would have the same reaction, which is that the drug had to be active against all three

pathogens. I think we can't minimize the non-typeable H. flu or the Moraxella, and there can be complications with those two. They are not totally trivial, but most of them get better on their own.

DR. CRAIG: I think again it is in the older children they are not very virulent, but the question is still in the ones under age two, whether they might be considerable there, and I think people have been able to--if you look at most of Ron Dagan's studies, where he has been able to identify things, he has got Haemophilus in those, and all of his patients are essentially under the age of two, so that I think those organisms can be pathogens in that age group, but it is in the older age group where we wonder about their significance.

DR. LEISSA: Can I just try to summarize?

DR. CRAIG: Brad, yes, you get the last word.

DR. LEISSA: I get to summarize.

DR. CRAIG: Yes.

DR. LEISSA: What I think I heard from the committee was to the issue of the clinical-only study, that although there are some concerns about the utility of the information that come from that study, there may still be value in that study also from the perspective of safety information, but that there was an encouragement to try to

"enrich" the population to some degree with regards to age, looking for children less than two; that in the microbiology study, that we all recognize that it is valuable from the perspective of getting more information about Strep pneumoniae especially in the non-susceptible, and that it was feasible, we would like to see taps at the three- to five-day on-therapy visit in failures, and ideally, also in those that were non-susceptible at baseline.

One thing we didn't really get into too much discussion--and I am not sure we could really come to any consensus--is the issue about how many Strep pneumoniae, whether in that number that is all susceptible Strep pneumoniae or only those relative to the ones that are non-susceptible, and that we should be working with industry to try to develop incentives for better data in terms of the issue about children who are failures, in terms of getting the re-taps done on those children.

DR. NORDEN: I need to sort of second what Barth said. I still don't see how we can really get efficacy data from the clinical-only trials. I don't see how we get efficacy data that is meaningful from the clinical-only trials, and so I respect the limitations that have been stated by our colleagues in industry, and I was in industry once and I know what they are, but bone biopsy is a painful

procedure also, and it is now required.

I don't think you can do efficacy studies in osteo without knowing the organism. I don't see how we can either in this, so I would like to get in as a strong second for what Barth has said.

DR. LEISSA: Would the Chair like to make a vote?

DR. CRAIG: You can make a vote, but let me make a comment before he makes a vote. Again, I think there are studies out there looking at children under the age of two, looking at bacteriologic failure and seeing what percentage of those have clinical failure, and you end up with, if you have 100 patients that are bacteriologic failure, about a third of those will also be clinical failures.

So, it is not as sensitive, but you can still pick up clinical failures if you are looking at that population that is less than two. It is the other population that we have diluted most of the studies with that I think make it difficult to pull up that out.

So, I am not convinced that you can't do a clinical-only study if you have it enriched with the patients that are there, that you would not be able to see a clinical difference. You would be able to do it with more sensitivity by doing bacteriologic data, and that is why I would try and do it in some way, as we mentioned, to try and

be an incentive for the industry to try and get that data at least in clinical failures and especially on any patients in the microbiologic study that have resistant organisms, because those are the ones that you really want to know how the drug works in.

Do you still want to vote, Brad?

DR. LEISSA: I see a shaking of the head.

The only other thing I think I heard was to the issue of the acute perforation. Dr. Reller was saying that he would believe that isolates of Strep pneumoniae and Strep pyogenes would be of value, but not so with Haemophilus and Moraxella.

Is that a correct summary of what you said?

DR. RELLER: I think it is a matter of relative importance, and it is also a question of duration. Once you get into the chronicity of the drainage, I mean you have got respiratory flora there, and it doesn't mean anything perhaps even with the others.

So, I think that one has to be very, very cautious in this slippery slope to a moist swab from the ear in saying anything about acute otitis media.

DR. LEISSA: Thank you.

DR. CRAIG: The next one is vulvovaginal candidiasis. The FDA presentation will be by Joseph

Winfield.

Vulvovaginal Candidiasis

FDA Presentation

DR. WINFIELD: Dr. Craig, Dr. Soper, other members of the committee, FDA colleagues, and invited guests: It is indeed a pleasure for me this afternoon to discuss with you vulvovaginal candidiasis.

The remainder of the afternoon will be shifting from the pediatric age group to adult females. I will be discussing vulvovaginal candidiasis, and the discussion following will be on bacterial vaginosis.

I am Joseph Winfield, a medical officer in the Office of Drug Evaluation IV, with a specialty in obstetrics and gynecology.

Actually, before I get into my presentation, I would like to share with you a short story that I heard when I was growing up in the South. It was about this minister that came into the community and was interested in getting parishioners to come to his church to worship with. So, he decided that he would drive around in the community to find people who had not been in the church with him and to entice them to come and worship with him.

So, he got into the automobile and he drove around the community and he came upon a farmer that was out plowing

in his field. So, he introduced himself. He said, "I am Reverend Jones, I am the new minister in the community in the church about three miles down the road."

So, he asked the farmer, "Are you a sinner?" The farmer replied, "No, I am a Johnson." He asked the farmer, "Are you lost?" The farmer replied, "No, I am over 60, I have been here about 60 years, and I know myself around pretty well." He said, "Are you ready for judgment day?" The farmer asked, "When is it?" The minister replied, "Well, it can be anytime, it could be today, it could be next month." The farmer replied, "Well, I am not sure, but whatever you do, don't tell my wife, because she will want to go on both days."

[Laughter.]

DR. WINFIELD: Hopefully, after my presentation this afternoon, I will be able to give you a better or communicate with you better than the farmer did with his minister.

[Slide.]

Historically, vulvovaginal candidiasis is an extremely common disease, 75 percent of all women will be infected at least once during their lifetime. This disease is second only to bacterial vaginosis in the causes of women visiting their doctor today.

Historically, in the 1950s, this disease was treated by the woman going to her physician and having him to apply Gentian Violet to the vagina or vulva. The problem with this was, of course, the severe staining that was encountered with this treatment, and also the long treatment that was necessary.

In the 1960s, the antimicrobial group polyenes were in development, of which nystatin is an example, and the problem with this treatment was, of course, the long duration of treatment that was necessary, from two to four weeks.

In the 1970s and 1980s, imidazoles were developed. Good examples are clotrimazole and miconazole. The advantage of these were that they were shorter treatment regimens, and they were more sensitive, the organisms were more sensitive to these drugs.

In the 1990s, we had the azoles expanded to the triazoles, and examples are terconazole and fluconazole, the advantage of these are that they are broader spectrum.

[Slide.]

Until 1990, all topical antifungals products approve for treatment of VVC were by prescription only. There were several sponsors in 1990 who suggested that their products should be able to go over the counter in order to

make them more available to the ailing female, so that she would not have to go to the physician necessarily to receive treatment.

The FDA took this under advisement, and in June of 1990, they conducted an advisory committee to look into the feasibility of the approved prescription products going over the counter. It was the recommendation of this committee that the approved 7-day imidazoles (clotrimazole and miconazole) be approved for over-the-counter use.

It was at this time, since 1990, we have had several durations of therapy of clotrimazole and miconazole, and others, that have gone over the counter for use.

[Slide.]

The intent of this VVC document is to provide guidance to sponsors regarding clinical trial design, evaluability criteria, statistical considerations, study endpoints as it relates to prescription drugs only.

Even though these criteria will be applicable to OTC products, consultation with the OTC Division will be necessary before we can publish similar guidelines.

[Slide.]

Study considerations. The Division recommends two statistically adequate and well-controlled multicenter trials be conducted that establish equivalence or

superiority to a topical approved seven-day antifungal.

The reason for the seven-day product is that we feel that we have seen through time that the best results in terms of cures are obtained with the seven-day products if all of the drug is used correctly.

Secondly, this will prevent the biocreep phenomenon.

[Slide.]

Further study considerations include randomization. We feel that all patients with a positive KOH should be randomized. That of blinding, double-blind trials are preferred, but in the situations where this is not possible, this is where you have a shorter duration of therapy compared to the seven-day, we would expect at least as a minimum, investigator blinded.

We are encouraging that all patients who receive drug should have a follow-up visit.

[Slide.]

Mycological considerations. In Phase II and Phase III clinical trials all fungal isolates recovered at entry and at follow-up should be identified to the species level, and in-vitro testing should be performed to determine the susceptibility of fungal isolates to the antifungal drug that is being studied.

[Slide.]

Inclusion criteria for our trials would include postmenarchal female with a clinical diagnosis of VVC based on the following vulvovaginal signs and/or symptoms: itching, burning, irritation, edema, erythema, excoriation.

[Slide.]

Each of the signs and symptoms mentioned before is to be scored between zero and 3, that is, 1 being mild, 2 moderate, and 3 severe.

To be evaluable or to enter into the study, we would like to require a minimum composite score of signs and symptoms of 2 and a positive KOH for hyphae/pseudohyphae, and culture be performed for Candida species.

[Slide.]

Exclusion criteria would include other infectious causes of vulvovaginitis, i.e., Chlamydia, bacteria vaginosis, herpes, HPV, et cetera.

For patients who receive any antifungal therapy 7 days prior to entering into the study, and in pregnant patients when appropriate.

[Slide.]

In terms of evaluability, the patients that participate in the study must have drug compliance. This depends on the treatment duration of the study arm, of

course. For the single and three-day therapies, these individuals should receive all drug; for seven-day therapies, they should receive the first three consecutive days of therapy.

[Slide.]

We then move to evaluation visits. The evaluation visits, this is a departure from what we were requiring in the past, and that we are eliminating or suggesting that we will eliminate the return early visit which used to occur usually seven to 10 days after the end of therapy.

We have an entry visit. At this visit we would expect a complete history and physical examination to be done, certainly including the pelvic examinations, and this would also rule out other infectious causes of vulvovaginitis. At this visit a KOH is performed. This is a screening test, and the fungal culture is done.

Also at this visit, diary cards would be provided to the patient and explained. The information regarding what is on the diary card or the information obtained on the diary card is included in the handout, but it includes information to be sure that the patient is compliant and when does the patient receive relief, and if there are other products that the patient used during the course of treatment.

We are suggesting, in place of the seven to 10-day return visit, that we have an interim phone call. The phone call is recommended, but it is not critical for evaluability. The purpose of the phone call is to ensure drug compliance, to determine early failures, and to assess adverse events.

[Slide.]

Further in the evaluation visit, we have the test-of-cure visit, and this is where the patient is evaluated in terms of cure or failure. We are recommending that the test-of-cure visit occur between days 21 and 30 after study entry.

The reason for this is we feel that at this time period, it is adequate to evaluate the performance of the drug and we can reduce the number of patients who do not return for follow-up. Previously, they were returning anywhere from 28 to 45 days post-therapy, and we feel that we can maintain many more patients by looking at the test-of-cure at 21 to 30 days.

At this visit, the test-of-cure visit, evaluation again of the signs and symptoms would occur, and speciation and susceptibility testing on all positive cultures. We would also have the investigator's assessment. This assessment is necessary because in some cases you have signs

that may be of minimal degree or you may have new signs that appear and that are not related to the disease, and if the investigator feels in his opinion that this patient would be a cure, then, he would assess the patient as such.

[Slide.]

We have outcome. At the study, we have what we call "outcome." We have clinical outcome, mycological outcome, and therapeutic outcome.

For the clinical outcome to be considered as a clinical cure, each of the entry sign and/or symptom that was scored as 1 or 2 should be zero at the test-of-cure visit. For each entry sign or symptom that has the score of 3 at entry, should have a score of zero or 1.

Any new sign or symptom should be assessed by the investigator as related or not related to VVC.

[Slide.]

If the patient does not meet the clinical cure criteria, then the outcome would be a clinical failure.

[Slide.]

For mycological outcome, the patient would have mycological eradication, and this would be negative culture for yeast at the test-of-cure visit or mycological persistence which would be a positive culture at the test-of-cure visit or earlier.

[Slide.]

The primary endpoint then would be the therapeutic outcome. This is somewhat new than what has been presented in other indications, and in the therapeutic outcome we are defining either cure or failure.

A cure is defined as the patient that would have both a clinical cure and mycological eradication at the test-of-cure of visit.

A failure then would be clinical failure or mycological persistence at anytime during the study period.

[Slide.]

For evaluability, patients can be evaluable either as a cure or as a failure. Assessment as a cure would have the clinical and mycological cures and eradications respectively, occurring between days 21 and 30.

No antifungal drug is to be given during the study period days 1 through 30 other than the drug in the tests, the test drug and the comparator drug.

As a failure, if the patient received antifungals between days 3 and 30, or in the investigator's opinion, the patient is a failure.

[Slide.]

For analytical considerations, we recommend analyses be performed on two populations: the

intent-to-treat population, this would be all randomized patients or all of those that had a positive KOH would be randomized, and the evaluable patients per protocol. Those would be meeting the protocols in terms of all of the inclusion and exclusion criteria.

[Slide.]

For statistical considerations, we are proposing that each study should be adequately powered to demonstrate therapeutic equivalence using a 95 percent confidence interval around the difference in the therapeutic cure rates of the test drug to the comparator for the per protocol evaluable population.

[Slide.]

In summary, what is different now that we are proposing than what we have done in the past?

No. 1. We are asking for two study visits instead of three.

No. 2. That the KOH be utilized as a screening tool only. Previously, we are requiring that the KOH would be counted as part of the mycological evaluation in terms of cure failure, and we feel that the culture is more appropriate to be considered in this evaluation.

No. 3. The test-of-cure window has been extended from 21 to 30 days, and we are only requiring the one visit

rather than the two.

No. 4. Speciation and susceptibility testing be done at both entry and the test-of-cure visit. This will give us some idea about the development of resistance or whether the patient failed because she developed a new species or we will learn more about the performance of the drug.

No. 5. Next, we have the consolidation of vulva and vaginal signs and symptoms. If you notice, enumerated on 6. Previously, we had as many as 18, and if you add that in terms of severity of the signs and symptoms, you can see that the numbers can get enormous and it can really be difficult to evaluate.

Finally, we are recommending that all patients who receive any drug would have a follow-up visit or remain in the study, so we can find out and tell something about the performance. Previously, if they did not take all of the drug, then, those patients were excluded from the study.

[Slide.]

This is the end of my presentation. I have no specific questions to the committee, and we will entertain any comments or questions or remarks from Dr. Soper.

Thank you.

DR. CRAIG: Any questions for clarification on

anything that was presented? Yes.

DR. BLACKWELDER: Of the two analyses you mentioned, when you said per protocol, did you mean to include only those patients who received enough drug as you defined it, were compliant?

DR. WINFIELD: That would be correct.

DR. CRAIG: Any other? Okay. Thank you very much, Dr. Winfield.

Dr. Soper will give the committee presentation.

Committee Presentation

DR. SOPER: I just essentially reviewed the remarks that were just made, and I just have the following comments. First of all, I think it is great that you have consolidated vulvovaginal symptoms to the six that you noted. I think there needs to be some guidance with respect to what mild, moderate, and severe is, and it would be helpful to quantitate that as opposed to leaving that up to our imagination of what mild, moderate, and severe is.

As far as inclusion criteria, on KOH, you noted the utility of seeing fungal elements, but you left out budding yeast or blastospores, and I think you can add that to inclusion criteria. Some patients will have only that finding when they present with acute vulvovaginal candidiasis.

I am supportive of the extended test-of-cure. I really never understood the utility of the early evaluation at the seven day mark. I think allowing at least two weeks after the completion of a seven-day course of therapy and maybe even considering liberalizing the window in which patients would be considered evaluable.

Right now it is 21 to 30 days, but remember that one of the reasons that patients become unevaluable would be if they couldn't get back to the office during that window. You might even consider liberalizing it to 45 days, 21 to 45 days.

That kind of is the double-edged sword given your change in what a therapeutic cure is because it gives the patient more chance to become mycologically positive, but it also gives the investigator additional chance to get the patient back in the office for evaluability, and it does allow us to assess early relapse, which I think is an important parameter in evaluating patients that have had an initial response to acute therapy, who then in a relatively short period of time, within a month, have recurrent symptoms.

The biggest change that probably industry is going to have a problem with is going to be with the therapeutic cure definition, and what you have proposed is that the

patient have both clinical resolution and mycological cure, and at first blush that seems to be incredibly reasonable, except that the difference in clinical response and mycological cure is pretty substantial, it is at least 20 percent, and so you are going to essentially dictate a lower efficacy rate for all future studies if you insist on therapeutic cure.

It doesn't really make any difference as long as we all know what the rules are because comparators are used, but I can see how maybe the efficacy if you were just using clinical cure would be 85 percent by one drug, and then if all of a sudden a new study is done, when rules are changed and their efficacy is 65 percent, then, all of a sudden the advertising says oh, our drug is so much better than yours, when really they are similar. That just needs to be addressed.

I think that there actually is some utility in changing the definition because you would like to see in the best of all possible worlds, patients that not only had a clinical response, but also were cleared of the pathogen, and that probably will predict those patients that are less likely to have early relapse, but the point remains is that the persistence of the pathogen doesn't necessarily predict that the patient is still going to have symptoms.

That is all my comments.

Committee Discussion

DR. CRAIG: Dr. Murray.

DR. MURRAY: I have a question on the requirement for susceptibility testing because I am not absolutely up to date on the susceptibility testing of fungi except I know there is an NCCLS Working Group, but I didn't know it was formalized enough to be an official group, plus some drugs may not even be on the official list, and certainly a new drug in investigation, one may not have any idea what its interaction with the test media, et cetera, would be.

So, I had some question about that and assumed that you probably had a fungal susceptibility consultant to advise you on this.

DR. LEISSA: I am not sure if Dr. Gosey is in the audience. Dr. Gosey, do you want to come up and address that comment? Dr. Gosey is a microbiologist with us in FDA.

DR. GOSEY: The NCCLS does have a subcommittee on antifungal susceptibility testing. Right now there are tentative procedures out there for the azoles, and what we recommend--I am on the committee--is that we use those same types of procedures as for fluconazole and intraconazole, and take it from there as to how the MICs change over time to eventually set the breakpoint for susceptibility

resistance.

DR. MURRAY: It seems a bit premature. What are you going to do with nystatin, has that even been evaluated, could that be a legitimate comparator? Would you require that then?

DR. GOSEY: No, we would use a comparator as an azole, as well. I am assuming we are testing mostly azoles at this point. We do have tentative break points for amphotericin, and I think we could go ahead and do that for nystatin. Again, those are in the polyene groups.

DR. MURRAY: I guess I just had concern because I think there are all still fairly preliminary, is that correct? You would know better than I.

DR. GOSEY: There are set breakpoints for intraconazole and fluconazole. As for the other azoles, we do not have set breakpoints at this point.

DR. CRAIG: Is that documented still at the tentative level or is it final?

DR. GOSEY: The document has gone through and it has been approved by the committee. It is early and we realize this, and it is going to be a growing process, but we do know that resistance does occur, and this is something that we need to start to get a handle on.

DR. CRAIG: Dr. Reller.

DR. RELLER: Two questions I have. What kind of specimen is one going to use to show eradication of the fungus, the yeast, and is that important in terms of clinical response? Is the recurrence owing to organisms that are resistant to the antifungal agent or recolonization in a milieux of ecologic imbalance from the gut?

DR. SOPER: The specimen is just a vaginal swab, and most recurrent vulvovaginal candidiasis is a relapse from identical strains that cause the initial episode. As a matter of fact, it is interesting to look at groups of patients that have recurrent disease and those that have responded to therapy in that they have similar culture positivity rates following therapy, but the patients that have so-called recurrent disease obviously have symptoms where the other patients do not, which is again one of the concerns of lumping the culture and clinical symptom data together with respect to therapeutic response.

It is felt that the etiology of recurrent disease is based, not only just on the presence of the microorganism, but also may have immunologic and allergic type, hypersensitivity type of parameters that lead to persistent symptoms or recurrent symptoms.

As far as the milieux, the bacteriological milieux, really, the microbiological milieux in the vagina

of a patient with yeast is pretty normal except for the overgrowth of yeast or maybe from the hypersensitivity or response of the host to the yeast as opposed to the next disease we are going to be talking about, which is a very complex alteration of vaginal flora, which is bacterial vaginosis.

DR. RELLER: I realize the needs for the clinical trials may be different from the clinical world, but I don't think most places are making the diagnosis of Candida or vulvovaginitis based on cultures of vaginal swabs. Are people missing the boat clinically or are we asking for data as an assessment of cure that is not germane to the issue?

DR. SOPER: I don't think you are missing the boat. I think it is important to insist on culture at least at entry level because it confirms microscopy. Again, as we will talk about the next part of the day, the ability of individuals to microscopically confirm the presence of microorganisms, you would expect would be relatively straightforward, but it is not.

It is a nice control, if you will, in yeast to have a positive culture confirm the KOH, and then bacterial vaginosis Gram's stain to confirm the diagnosis.

DR. RELLER: Are there yeast or even Candida species other than albicans that cause this entity?

DR. SOPER: Yes. The most common species, of course, is going to be Candida albicans in causing acute vulvovaginal candidiasis, but non-albicans species, like glabrata, krusei can cause acute, and even Baker's yeast can cause acute vulvovaginal candidiasis.

DR. RELLER: So, the culture initially is to confirm how good the KOH preparation was for recognizing yeast?

DR. SOPER: Yes, as well as to maybe identify those species that might be more resistant to triazoles or imidazoles that are being tested these days. In other words, it is more likely that albicans species are going to be responsive to traditional therapies used today than non-albicans species.

DR. RELLER: But the presence of any quantity of yeast post-therapy is a good marker for efficacy?

DR. SOPER: No, it doesn't necessarily relate to clinical resolution of symptoms. I think most of the time what you would find is patients that were without symptoms would be culture-negative, but a substantial proportion of time, I am talking 20, 25 percent, they may be persistently culture-positive.

DR. RELLER: My thought is that this is just too demanding an endpoint that is not related to the clinical

reality, but I could be missing something. This gets into a very--it is difficult.

DR. SOPER: The way data was reported, it used to be clinical cure, which were essentially the resolution of clinical symptoms, and mycological cure, and they were reported separately, and essentially, clinicians look at both.

If you had an agent that was associated with an accepted clinical cure, you could count on the mycological cure being substantially less, again on the order of around 20 percent. If I was looking at new agents and I saw that the clinical cure rates were similar, but their mycological cure rates were inferior, I would probably use the older agent. So, I think there is utility despite the limitations of culture as a test-of-cure in reporting that.

It is one of the reasons why I don't have much of a problem with the recommendation that the therapeutic cure term be used as a combination of both clinical resolution and mycological cure. It just is important to make sure everybody understands what that means and what the difference from previous literature is.

If you are going to quote, say, efficacy rate of a certain imidazole as 85 percent, that that means in the old literature clinical resolution of symptoms. If you are

going to say the therapeutic efficacy in a new study is 65 percent, the clinical resolution may still be 85 percent because of the mycological discrepancy, you are going to report your data a little bit differently.

DR. RELLER: If there are clear-cut data that with a given there is an increased likelihood of recurrence of the clinical entity based on persistence of the organism after clinical resolution, you know, at some early time period, recognizing that the longer one goes and becomes again difficult to separate out these persons who are persistent carriers, and some data, I think reasonable data, that it may be necessary to greatly diminish the numbers in the gut if one is going to prevent recurrent vaginal candidiasis, but making the analogy, I mean with Helicobacter pylori, if one eradicates the organism, it is a marker for preventing recurrence of some clinical entity.

In contrast, I know of no data that really solidifies that, and it is not standard practice, and I don't think it should be, to if you treat someone with a drug that is efficacious for Group A streptococcal pharyngitis, you don't do throat cultures to show that the organism is gone, because if one does, you know, in someone who clinically responds and everything else is going fine, I mean you can find an organism, but that doesn't mean you

keep treating it over and over and over again.

I think we need to be sure why we are doing this, that's all.

DR. SOPER: I agree, and Bill actually is exactly right. In true clinical practice, we don't reculture. Actually, we don't culture patients to make the diagnosis in the first place, only those patients that have chronic recurrent disease to identify for the most part non-albicans species, but in a clinical trial in which you are studying a new agent, I think a test-of-cure culture is appropriate.

DR. CRAIG: And is done in pharyngitis for clinical trials.

Dr. Leissa.

DR. LEISSA: I just want to give some historical perspective in that this issue of the therapeutic outcome, which is a composite of the clinical of the mycological, has been used by the Division for at least 15 years or something like that, so this isn't a new phenomenon.

The major change here is the issue about doing the microbiologic susceptibility testing. Before, indeed, the issue when it came to mycologic evaluation, all cultures that I know of in most of the recent studies use what was called a "biggie" culture, which most people don't probably even recall, and we are recommending in the guidance that it

actually use Sabouraud's dextrose agar to the issue of mycologic.

So, therapeutic outcome has been around for a long time. It is just to the issue that Dr. Reller is raising, which is should these be inextricably linked to the overall issue of efficacy or whether they should be disassociated, and you look at clinical and you look at mycological separate, and look to see what the overall, whether again the same findings are coming from both populations, because the issue of therapeutic outcome is unique actually of all the indications discussed to both vaginal candidiasis and what will be discussed later, bacterial vaginosis.

DR. WINFIELD: I would like to say that in the clinical trials that have been conducted--and we have looked at it both ways--we have looked at the clinical cure and the mycological eradication, and what has happened, like in the earlier visit, what you would have, you may have what you could call a clinical cure, but then at the later visit, you would actually have the recurrence of clinical signs and symptoms.

If you look at a lot of those instances, what would happen is that the organism was not eradicated even though the signs were abated. So, if you really look at clinical trials, what happens with the therapeutic outcome

or the therapeutic cure rate is almost identical to your mycological cure rate in terms of the rates.

They may vary something less than 2 percent, but your clinical cure rate is usually going to be much higher, as Dr. Soper said, but the mycological, we feel is a true measure of what really happens with the organism.

If you have a mycological cure, most of those patients will not end up with a recurrence of disease within the evaluation period.

DR. CRAIG: I am still a little confused of why we need to put the two together when for virtually all other indications we do them separately.

DR. DAVIS: I would just make a comment concerning that. I think that if you have the two entities separate, I think there is some tendency for industry to just pick what is most favorable and to their advantage, and I think since these really are clinical trials for the approval of a drug product, I think you do need and should have your strictest criteria, and the strictest criteria, in our opinion, is to combine the two and have a therapeutic outcome that is based on both the clinical response and the microbiology data.

Again, these are comparative studies, so you do have a comparator arm and then your study arm, so that should the efficacy overall be, let's say, lower than it may

have been 10 years ago, at least it is comparative, so that you always have the comparator arm and the study arm, and both may show lower efficacy, but I think it is just a more well-defined and stricter endpoint and analysis for approval of the drug.

DR. CRAIG: I guess the question I have, where I get a little concerned, if somebody is a little bit more vigorous in taking their specimen, are they going to have a greater chance of getting a few organisms that are still left around to have a positive culture even though from a clinical cure point of view, there is no difference, or is it going to be roughly the same no matter what kind of populations, they are always 20 percent that you have.

If there is something else that can vary that percentage, that is not necessarily related to the drug, i.e., how you do the specimen, what kind of swab you use, maybe some of the fungi stick better to than others, then, I question using that as a criteria, because I think one of the things we frequently want to do with studies, that frequently happens, is we also like to even be able to look at one study and compare it with other studies even if they are using not necessarily the same comparator.

That is one of the goals, I think, is to try and make the studies pretty much the same, and if there is

something else that comes out, I would feel much better being able to look at the clinical cure and knowing what the clinical cure is and be able to go across that, and then look at the microbiologic and say maybe there is the area where there is some difference than to put the two together and sort of lose the separate evaluations that one would get.

DR. WINFIELD: The problem you will have, though, with your clinical cure if you just use that, or which one would you use, would you use the clinical cure or the mycological cure to determine the performance of the drug? As I mentioned earlier, if you were to follow the patients, and if you look at them clinically, after they have finished the drug, clinically, they may consider themselves as cures, but then if you look at them four weeks later, a good percentage of these patients, 10 to 15 percent, the symptoms have recurred.

On the other hand, if you look at the mycological eradication, at the early visit and at the long-term follow-up, those who have negative cultures at the early visit, say, 7 to 10 days after therapy, as well as 28 to 30 days or 35 days after therapy, they still were negative in terms of culture.

So, I really don't have a problem in terms of

eliminating the therapeutic outcome provided we would use the mycological cure as the primary endpoint in terms of evaluating a drug.

DR. SOPER: I think this is more semantics than anything. I would support using your therapeutic outcome, but I think the data can be reported so that you can compare old data with new data, and that is that clearly, industry and investigators are going to want to share both clinical outcome, as well as mycological outcome in separate presentation, and then combine it as a therapeutic outcome when they present the data.

DR. DAVIS: I would like to add I just approximately three months ago did finish a major NDA for an antifungal, and, in fact, both the sponsor's analysis and mine really did look at clinical outcome, and that is one set of data, and then the microbiological, but then I looked at those patients who had both the clinical and the mycological outcome, so the clinical efficacy was--I am going to make this up a little bit--let's say 80 percent, mycological was about 80 percent, but, in fact, if you then found patients who had both entities, your efficacy rate was about 65 percent.

So, if a sponsor, in fact, wants to go back to that study and say, but our clinical outcome or cure rate

was 80 percent, they actually can do it from my analysis, as well as their own analysis.

DR. CRAIG: Just to see if I can explain a little better, if you have primarily a fungistatic agent and where you are sort of dependent, you are just slowing the growth of the organism, you sort of hope that may affect its ability to adhere, and so it will sort of just be washed out and disappear.

You might find something entirely different with the fungicidal agent that actually kills the organism, so I could see the scenario where clinically, the two were equivalent as far as treating the disease, but in terms of having that 20 percent that stays positive, the fungicidal agent did better.

What you would then be saying is that it is necessarily a better agent, which I am not sure for treating the disease that it necessarily is the better agent, and that is where I am trying to explain why I don't see combining it together. In my mind, it is more the clinical effect of the drug that would be the important aspect.

DR. SOPER: I think what will happen is that by extending the test-of-cure, that the symptoms will catch up with the culture, and that the early test-of-cure and the mycological discrepancies is where the discrepancy occurred,

and that by eliminating that early visit, you really get a better sense of what is going to be longer lasting, essentially the outcome you really want to evaluate, the longer lasting resolution of symptoms.

DR. CRAIG: Any comments from industry?

DR. WITTES: Could I make a couple comments?

DR. CRAIG: Yes.

DR. WITTES: There is two issues that I would like to bring up. One is what I think is the stringency of the clinical outcome. The way I read it, somebody could come in with mild itching and mild irritation, a score of 1 of each, giving 2, reducing to zero as a cure.

Somebody else could come in with six 3's, reducing five of them to zero, and one of them to 2, and that is not a cure.

I think that there needs to be some kind of consideration about the consequence of the scoring system, which I think can lead to this kind of inconsistency. That is one issue, and let me bring up one other one which I am sure the three statisticians know that I would bring up, the issue about the per protocol analysis.

I think that the way--and you all have to know that reflexively I would react against it--but in this case, I think there is actually another kind of potentially

illogical situation that could occur.

A seven-day person is supposed to receive the first three consecutive days of treatment. You can imagine somebody coming in, be treated at day 1 and 2, skipping 3, and then the rest of the four days be in there, and she is excluded. Somebody who doesn't start on day 1, starts on day 2, and continues.

So, I think that if there is going to be, if the primary analysis--and I am using the word because it is here--if the primary analysis is going to be a per protocol analysis, I think, first of all, you need to think about really whether that should be the primary, but if it is, I think it needs to be thought--the various scenarios about complying and not complying have to be worked through very carefully, so that you don't run into illogical inconsistences and who is in and who is out.

DR. WINFIELD: Did you say that they could take the first two days, skip a day, and then go the next day and take it and be included?

DR. WITTES: No, excluded.

DR. WINFIELD: Yes, they would be excluded.

DR. WITTES: Right.

DR. WINFIELD: The reason for that is that we feel that you have to have, before you could be considered a

failure, you would have to have at least three days of therapy before we would consider you as a failure, and therefore, the flip side of that is before you consider a cure, you would have to have at least three consecutive days of medication.

DR. WITTES: But, you see, I would call that three consecutive days.

DR. WINFIELD: Pardon me? Three consecutive days by skipping a day?

DR. WITTES: If you have 1, 2, skip, 4, 5, 6, 7, that looks to me--

DR. WINFIELD: No, the first three consecutive days--okay, the first three, right.

DR. WITTES: I understand.

DR. WINFIELD: The other issue that you talked about in terms of including patients--and this is another advantage of being able to combine the two in terms of therapeutic outcome or evaluating them therapeutically--because what would happen clinically, the patient may be able to enter.

In addition that, though, she also has to have a positive culture in order to be classified or to be evaluable. So, you are not only looking at the clinical signs and symptoms in terms of cure or failure, but you are

also looking at the mycological aspect before you will classify that patient as a cure or failure.

So, we would need both of them, we would feel, in order to give an adequate evaluation of the patient.

The other thing, it is rare, it is very rare for an individual to come in with all severe signs and symptoms. We would have to look at that patient on an individual basis, but that is a very rare occurrence.

DR. WITTES: I understand that, and I took obviously for emphasis, I took the most extreme, the most extreme/least severe, and the most extreme/most severe, but I think that again, to me it exemplifies what can happen with a very complex scoring system, which this is, admittedly much less complicated that previous ones, but nonetheless, it is a six-item scale, each of which has four points, and that is a complicated system.

And just to make sure that the definition of cure, I mean the fear would be that one drug that really makes a dramatic difference in the way women feel, it doesn't show up because in order to count as a cure, you have to do extremely well when somebody is in real pain and when the symptoms are very varied, and that is all I am asking for, some kind of thinking about how to deal with the most severe cases.

DR. DAVIS: Let me make a quick comment about that, and Dr. Winfield can correct me if I am wrong, but in the past, the criteria was that all of the signs and symptoms had to disappear. Our new recommendations are, in fact, that we would allow somebody with a severe itching or severe burning could go from severe to mild, and still be considered a cure as opposed to having to go from the severe to no signs or symptoms.

So, that is a change in the guidance recommendations, and it is true, a person could go from the score of 18, which would be the six-symptom score of 3, to a total score of 6, meaning mild for all six symptoms, and still be considered a cure, but that is, in fact, quite a change from I believe has been done in the past, and maybe some of the advisory committee members actually have a comment or concern about that.

DR. SOPER: What guidance are you going to give industry concerning the use of an additional, say, topical steroid in addition to the antifungal? It is not uncommon for patients to receive antifungal therapy and topical steroids for their vulvitis.

DR. WINFIELD: Those patients would be excluded. In the document, we have that they cannot use that drug, any topical product.

The other thing that is going to help us a lot is in the investigator's opinion, whatever the signs that the patient may present with, it is in investigator's opinion as to whether or not this patient is considered a cure.

Based on that, it will be whether or not he feels that she will need additional treatment. So, in some of these issues, we are not going to go strictly by scores in terms of number, we are also asking the investigator to give his assessment or her assessment as to whether or not the patient needs additional treatment.

DR. CRAIG: Barbara.

DR. MURRAY: I certainly would want to agree with you. One comment about the consecutiveness of the days, because as it is written here, it looks like if you took it on days 1, 2, 4, 5, and 7, you would be excluded, and I am not sure if I see a biological reason for that.

Getting back to your point of reason for stopping, if somebody felt very, very, very much better, that could be a reason for stopping, so it seems the consecutive days aspect seems a bit arbitrary unless there is a biological basis for it.

DR. CRAIG: Let's say on this one particular item.

DR. LEISSA: Yes, to that point. One of the issues that we wrestled with--and this goes across

indications--is when we were looking at the studies, to try to determine whether or not the study supports the requested indication and in the dosage administration, which is the number of days that the company says our drug works in this indication for seven days, in the past, there was a bias that a number of reviewers would say, well, if you are looking for seven days, that evaluable population would be patients who took six, seven, and eight days, because that is around the requested time period.

So, one of the things that we thought was that if we are thinking patients can be failures as early as three days of therapy, what do we do with the patients that are four, five, and six, because in the old realm we might make those non-evaluable because they hadn't had 80 to 120 percent of the drug.

So, what we were thinking here was that, well, let's not throw out those patients, let's not throw out the four, five, and six day patients, but in the interests of doing that, let's also not bias ourselves by saying we will accept patients on their outcome depending on how many days of therapy they took.

So, it was kind of an arbitrary decision, so that if you had a patient who took days 1, 2, and then they went 4, 5, 6, but they were a failure, well, what if they had

taken the first three days, would they have had a different outcome. Part of it, like I said, was an arbitrary way of trying to be at least clear about how those patients should be counted.

DR. SOPER: That is actually unusual to see that early type of noncompliance because it takes an average of two, two and a half days for patients to become asymptomatic whether they are treated with a single dose or multi-dose regimens. So, you are going to see noncompliance I think after four days, five days.

DR. CRAIG: Yes.

DR. BLACKWELDER: For both the issues that Janet brought up earlier, I would like to support her comments, and I think the last couple of minutes have pointed up how complicated defining compliance could be, and if the statistical objective is to show that the new drug is similar to another one, it seems to me that it might be worth considering rather than to define one called per protocol or primary analysis, that there might need to be a variety of analyses before you are comfortable with saying the new drug is as good as the other one.

That is one comment. Can I make one on the clinical scoring?

DR. CRAIG: Sure.

DR. BLACKWELDER: I would suggest that at least you consider something like maybe as a secondary analysis, defining how much change the woman saw, maybe based on what her score was to start with, that is, if she started with 2, she could only get better by 2, if she started with a 10 or an 8 or a 6, she could get better by a lot more.

I think there were a couple of comments before supporting looking separately at the mycologic and clinical outcomes, and it seems to me that that would be a wise thing to do, as well.

DR. CRAIG: Yes.

DR. JENNINGS: Good afternoon. Cherylisa Jennings from [Pharmacokinetics] Laboratories.

In looking at the guidelines on page 4, under Signs and Symptoms, you specifically pull out vaginal discharge, that that should not be used as one of the signs and symptoms. We would just like to know why that was.

Also, you didn't address in your presentation pap smears. That was one of the things that was recommended also on page 4, and then just to address the use of condoms is trials, the criteria in the guide excludes patients using condoms.

DR. WINFIELD: We deliberately eliminated discharge because there is a physiologic discharge, and we

have gotten into a lot of problem in terms of what is physiologic and what is pathologic. So, we thought it was even though you would have a discharge to enter the trial in terms of what they are seeing to suspect that the woman has VVC, we feel that using a discharge is a parameter and evaluation is not appropriate.

We are asking that pap smears be done. This is part of the history and physical, and it is just part of good clinical practice, and we are also recommending that patients who would have an abnormal pap smear or carcinoma in situ, et cetera, that they would be excluded from the study simply because, for one, that we would prefer that that woman would have that condition taken care of, and it may have some bearing on the results.

In terms of condoms, this is part of the diary in terms of what other devices they have used. We are not really excluding those patients, but if they have used other devices, we will be looking at those patients and what effect that the condom may have had on the product.

DR. CRAIG: Yes.

DR. FOX: Barry Fox from Bristol Myers. A couple of comments and then one question for the agency.

I wanted to first second the comments on the floor regarding the scoring system and the improved categories,

and potentially the person having one number that was the same, and then potentially calling that person a failure. I assume that we have to entertain other types of vaginal infections which will be discussed next, that may even crop up in the four weeks after the patient is treated.

Secondly, the agency has encouraged in oral thrush studies doing actual quantitative cultures for yeast to look for potential enrollment and endpoints, and I think it is a little bit fraught with potential difficulties to look at a plus/minus system of just a positive or negative culture in the vaginal area. It is certainly not feasible to do quantitative cultures in the vaginal area, but this just raises the concern that a black and white or a positive or negative culture may not be a proper endpoint.

The third issue is did the agency consider issues of prior treatment with topical antifungals, since it has been made so easy for women to get the agents over the counter? Patients may not come to their physicians or the investigators immediately, and it may be prudent to consider allowing a dose or two of topical antifungal therapy provided that the KOH is positive, and also provided that the culture does grow from the initial specimens.

DR. WINFIELD: We felt that any individual who had received antifungal therapy seven days prior to entry, they

should be excluded, because it may have an effect on the signs, although she may have a positive KOH, she probably wouldn't have the appropriate clinical signs to include her in the study, so we excluded those, and it also may have an effect on the KOH.

So, we feel that a week prior to that, it would be appropriate to exclude those patients.

DR. CRAIG: Any other comments, suggestions?

DR. LEISSA: I just wanted to go over a point that Dr. Soper made in his comments to us. The issue had to do about inclusion criteria, the issue about positive KOH for the hyphae/pseudohyphae, and you had recommended the addition of including budding yeast, I believe.

We had some discussion internally as we were developing this document to that, and I would invite Dr. Gosey if she would like to make any comments about the issue of why we would not potentially want to include patients who just had budding yeast.

DR. GOSEY: I think your original question about adding the budding yeast, as Brad said, we discussed it quite a bit. Typically, if a female has vulvovaginal candidiasis, and it is due to Candida albicans, the invasive form is the hyphae or pseudohyphae that is seen.

The only time that we would typically see budding

yeastlike cells causing infection would be from the organism Candida glabrata, and I have to agree that would be the only time that we would tend to see that.

Typically, if Candida is in the vaginal canal and it is not causing infection, you may see budding yeastlike cells here and there. We had even talked about semi-quantitating the KOHs to get a better idea as to what forms were present.

DR. SOPER: Are you trying to exclude those microorganisms?

DR. GOSEY: No, not at all. That is why we are having the culture there.

DR. SOPER: So, what is the problem with going ahead and including those patients in the protocol?

DR. GOSEY: I don't see a reason why we have to exclude them. My main reason is to have the types of fungal elements described in KOH. I just don't want it positive or negative. I want to know whether there were budding yeastlike cells seen. I would really like to know semi-quantitatively rare budding yeastlike cells, numerous pseudohyphae seen, something like that. That is personally what I would prefer to see.

DR. SOPER: Maybe I am misunderstanding what we are talking about, but my issue was that if you see a

patient, she has yeast symptoms, she has budding yeast on wet mount, she should be included in the protocol. That is my only point. And it was not included in the inclusion criteria, so that patient, if she didn't have pseudohyphae or hyphae, would be excluded.

DR. LEISSA: Right, and I think the concern if I am recalling the internal discussions correctly, was that if you had, say, C. albicans and there was budding yeast, that there would be increased concern that that actually might represent colonization.

DR. SOPER: Oh, I see, instead of disease.

DR. LEISSA: Right.

DR. SOPER: I would disagree with that.

DR. LEISSA: The last thing I just wanted to revisit, I think I heard from the members of the advisory committee concerns about the therapeutic outcome from the perspective of whether it really makes sense, that it seems like there is the clinical population, the mycological population, and to look at those potentially separately, and I think Dr. Soper raised the issue of whether we are talking about semantics here.

The only thing I think that has import to the companies is when they are actually doing their sample size calculation, and thinking about to be of the "per protocol"

evaluable population, that those patients were evaluable in the therapeutic population, if you are not evaluable in either the clinical or the mycologic, that affects your evaluability to the therapeutic, and that has import, I think, to the industry.

But is it fair to say that the people in general are concerned about our using this therapeutic outcome for this indication?

DR. CRAIG: I would say yes. The question is of combining the two or looking at them separately.

DR. LEISSA: Looking at two separate analyses versus having a primary analysis which includes both.

DR. CRAIG: Dr. Reller says separately. So, we have got two votes for separate.

DR. NORDEN: Separately.

DR. CRAIG: Three.

DR. WINFIELD: May I just ask the committee, though, if you have a separate analysis, what is going to happen. Say that you have a drug that clinically, you are getting an 80 percent cure rate, and mycologically, you are getting a 50 or 55 percent cure rate, how would you determine the performance of that drug, or whether or not you think that drug is appropriate for the treatment of the disease.

DR. CRAIG: If you are telling me that it has got a 55 percent microbiologic cure rate, and that there is no difference and you follow the patients on later, that it doesn't necessarily reflect relapse, I don't know. I mean I am still happy with the 80 percent cure. That is what I would be looking at.

But if you tell me that the 55 percent microbiologic is going to reflect what one is going to see farther down the line, and there is going to be a significance of having recurrent infections, then, the label should say that. It has been the whole question with chronic bronchitis, do certain drugs in chronic bronchitis delay the time before one gets into other exacerbations. We don't necessarily put that into the initial clinical trials for doing that.

I think whether it treats the infection or whether it prevents later recurrence of the infection are two issues, and I would hope then that the data would give that. If you are trying to wrap that all into one, then, I think you are always going to have a fungicidal agent has the potential always to give better overall results and be the primary agent that one is going to use as compared to a fungistatic because it has a greater change of keeping the organism around and having persistence and a later relapse.

DR. SOPER: I think the issue is that we have changed two things at one time, and that is that you have changed the test-of-cure to an extended period of time, and you have changed the definition, so they are both going to catch up with each other essentially and be compatible, but we don't know that with certainty at this point, I guess.

I guess you might if you really dissect the data like you have probably done, where the early evaluation was associated with this discrepancy, but the late evaluation was more related to culture positivity.

DR. WINFIELD: The only problem with that is when you stretch it out too far, like going to 40, 45 days, you lose so many patients, and then that makes the trial, that has to be so many patients in order to be evaluable, it makes it a huge trial.

So, this was an effort, that if we were to keep the therapeutic at this short time frame, I think we will be able to keep many more patients in the trial, and we can get an appropriate index as to what is really happening with the drug.

DR. CRAIG: Dr. Murray.

DR. MURRAY: Were you going to address this particular issue, because I was going to switch gears for a minute?

DR. IMMALLIS: Yes, I was.

DR. CRAIG: Go ahead.

DR. IMMALLIS: David [Immalis] with Advanced Care Products.

I just want to reiterate that I still thing that the therapeutic outcome is a valuable outcome both in reference to comparing data that we develop on new products to historical controls, but also because the clinical cure is at this point in time a subjective measure principally. There are no validated measures, and in any other therapeutic area that I know of, if you are going to measure subjective outcome, you would first develop a validated instrument that you know is reproducible and predictive.

I don't believe that that is the case here now, so that to put too much emphasis on a subjective variable measure, that probably varies site by site for the sign measures, and patient by patient for the symptom measures, to put too much emphasis on that I think is going a little too far at this point in time.

I think there are opportunities to improve that, but I don't think we should go there now. I do think the culture remains an objective measure, and I think it should be given the appropriate weight.

DR. ALBRECHT: I guess the other question I would

like to raise--and this is just sort of thinking it through in the context of some of the other indications we have discussed-do we know that women with vulvovaginal candidiasis may not actually have some degree of resolution of signs and symptoms spontaneously over the course of weeks, and so that looking simply at the clinical change over time, is there not some parallel like in otitis, that some women simply will get better with time or change in the cycle, the monthly hormonal cycles, so that again, to sort of I guess reinforce what was said, having the mycology helps you believe your clinical picture.

DR. LEISSA: The other issue, I just wanted to make sure that we are hearing from advisory committee, deals again with the microbiology, which is although some companies in the past for this indication have on their own done microbiology and done susceptibility testing, the one difference here is, is requesting or proposing using Sabouraud's dextrose agar, and secondly, doing susceptibility testing on all baseline isolates and if there is an isolate that grows later. This is a change, and it is a change to industry from an expense issue.

So, there truly is value that we should be proposing this for all future studies in this indication.

DR. MURRAY: I was actually going to ask another

question, how are the data going to be used? Are the results of susceptibility testing going to be used to exclude somebody, is that analysis going to be stratified based on the susceptibility testing, or are we just collecting this data, just going to be collected to have it, because the comment was made to see if they changed later on, later on when, in that patient or in future studies, I mean the companies, if they are going to do susceptibility testing at all, may find data from isolates from other sites. So, how is that going to be used? I think that ought to be considered.

DR. LEISSA: Any other comments from the advisory committee members about the issue of susceptibility testing?

DR. RELLER: I would only reinforce Dr. Murray's comments, but also I think it would be important in any susceptibility testing that the agency ask of a sponsor that it be restricted to drugs for which interpretive criteria were published by the NCCLS.

This is a very difficult area, and they spend an enormous amount of time trying to bring some science to it, and if one strayed, there is no point in trying to reinvent that, particularly with limited numbers.

I think one of the reasons for being cautious about understanding why to get this is because some agents,

there are not going to be validated interpretive criteria, so is it fair to ask that this--say this must be done when some drugs are going to have to do it, and some not, simply because it has been worked out for some agents, and not for other agents. It seems an important issue.

I think analyzing these data for clinical--on another point--clinically and mycologically, you know, ultimately, the agency decides whether a drug ought be approved and one wants to have guidance, but I think that the capacity to delineate what the drug is doing in terms of eradication of the organism, as well as clinical, is important to maintain.

They, too, should be together, and it would be great if they are together, but I could conceive of a situation where two drugs look exactly alike, 60 percent efficacious by the combined criteria, because they both eradicate the organism similarly, but one drug is associated with 85 percent resolution of symptoms, going from a high score to low, or ablation of any score.

I think the patient and/or physician would want to know the drug that has the clinical efficacy associated with it, when it is not associated with any more relapses regardless of what the swab shows on a Sabouraud's plate.

DR. CRAIG: For urinary tract infections, we don't

require complete sterilization. We want them below a certain number, but we don't go down to complete sterilization just because the techniques are such that that is hard to do. So, I still worry about a swab getting some Candida out at 21 days down the line.

DR. LEISSA: But the patient is fine.

DR. CRAIG: Yes.

DR. ALBRECHT: At the risk of raising this question again, is there any role to start discussing quantifying those swab cultures?

DR. CRAIG: I don't think the science is there yet.

Anything else? Yes.

DR. GOLDBERGER: Dr. Soper, do you think that the vehicle for whatever else is in the product would have any clinical effect, although probably not a microbiologic effect, if it is administered for several days unless it is a placebo effect?

DR. SOPER: I think the vehicle could have both a placebo effect, as well as an emollient effect, which can be somewhat soothing, and actually even, in contradistinction to that, because many of these vehicles have propylene glycol in them, it can have an exacerbating effect and cause burning and worsening of symptoms.

DR. GOLDBERGER: That is another reason obviously to be look at the microbiology data in combination, which you obviously don't want to be looking at this nonspecific effect.

DR. CRAIG: Let's take our break.

[Recess.]

DR. CRAIG: Bacterial vaginosis. The FDA presentation will be given by Dr. Dan Davis.

Bacterial Vaginosis

FDA Presentation

DR. DAVIS: Good afternoon. We will be presenting our new guidance document recommendation for bacterial vaginosis. My name is Dan Davis. I am a medical officer in the ODE IV. Prior to joining the FDA, I was in clinical practice for approximately 20 years and have been with the FDA for about a year and a half.

It is my pleasure to be here this afternoon and present our guidelines. My talk will essentially cover the highlights of the document, not absolutely every point but the highlights, and it will be similar to the format that Dr. Winfield used because they really do use the same template.

[Slide.]

I want to start with a little bit of current

clinical background about the significance and prevalence of bacterial vaginosis. It is the cause of approximately 40 to 50 percent of all vaginal infections, and, in fact, is the most common cause of vaginitis in women of childbearing age in the United States.

Risk factors for contracting BV include prior pregnancy, sexual activity, and especially in new sexual partners or having more than one sexual partner, and it is associated with recent antibiotic use and often is found with Trichomonas infection.

It is associated with several other infections: recurrent urinary tract infection, pelvic inflammatory disease, and postoperative Gyn and postpartum infections, more specifically, postpartum endometritis has been clearly associated with BV, postabortion infections, and posthysterectomy cuff cellulitis.

[Slide.]

It is also relatively common in pregnancy. This is becoming more and more of an issue throughout the United States. Approximately 15 to 20 percent of pregnant women in the United States have BV at some time during the pregnancy, and this is of significance in that it is associated with the complications of chorioamnionitis, of preterm labor, prematurity, and premature ruptured membranes.

Currently, there are two FDA-approved regimens for treatment of BV in pregnancy, and these are metronidazole and clindamycin.

[Slide.]

The history of therapy or treatment for BV goes back to the 1960s and 1970s when very poor efficacy was found with the use of the triple-sulfa creams, with the use of erythromycin and tetracycline antibiotics. Other studies looked at povidone-iodine in the form of either of a gel or an insert or a douche, again with poor efficacy, and other vaginal douche products were also studied.

In the 1970s, we began to see what were considered moderate cures with ampicillin and amoxicillin, and by "moderate," if you take several studies lumped together, the cure rates were approximately in the 60 to 65 percent range.

[Slide.]

In the 1980s, we began to see good studies, large studies with metronidazole and clindamycin. These were essentially used off-label, and studies included both oral therapy and topical therapy.

By the 1990s, we started getting FDA approval of regimens, which currently are approved for three, five, and seven-day therapies, both oral and topical therapy of both the metronidazole and clindamycin, and all of these products

are currently available for BV by prescription only.

[Slide.]

The objectives of our document are to basically present guidelines concerning clinical trial design, evaluability criteria, endpoints, and statistical considerations.

[Slide.]

The study considerations are quite similar to that for vulvovaginitis. We are recommending statistically adequate and well-controlled trials, that will be multi-center, establishing either equivalence or superiority to a comparator drug product.

The comparator should be a drug that is considered to be the standard of care and that is already approved by the FDA for the treatment of BV.

I do for a moment want to go back to the very first bullet. One of the questions that I will raise at the end of my talk is, in fact, the feasibility or advisability of actually having placebo-controlled clinical trials as opposed to active-controlled trials, but that will come up at the end of my talk.

[Slide.]

Further study considerations are that randomization would be all patients meeting the inclusion

and exclusion criteria.

The follow-up, this is very simple, just try to maximize--and this point has been made repeatedly throughout the day, and I would basically like to make it again, and the reason for that is that in the past, when the studies have been submitted to the FDA, many of the patients have been found to be non-evaluable, primarily because of minor protocol violations or because they were not followed to the end of the study because of some compliance problem, it was a minor protocol problem, and we really strongly encourage the industry to maximize the follow-up and allow much more analysis of the data that would come in under those circumstances.

Blinding, the trial should be at least investigator-blinded, and obviously double-blinded where possible. We would recommend that both arms of the study, the comparator and the study arm, have the same routes of administration. The reason for mentioning this is, as I said earlier, there are both oral treatments for BV and then the topical vaginal treatments, but we do recommend the same route of administration.

By no intravaginal placebo, I specifically am referring to those studies where there is a intravaginal product being used and where the two study arms differ in

their length. If the study drug is only three days, let's say, for treatment, and the comparator is a seven-day treatment, we are actually, in fact, recommending that you do not use placebo for days 4 through 7 in the shorter arm, and our reasoning for that is that we feel that the use of an intravaginal placebo may alter the pH of the vagina or the characteristics of the epithelial cells, or may allow for displacement of the study drug or dilution of the study drug, and all of these factors could then affect the outcome or efficacy of the study product.

[Slide.]

Microbiological considerations. This is one of those rare circumstances where, in fact, cultures are not recommended, and that is very simply because there is no clearly established pathogen as the etiologic agent of bacterial vaginosis. However, cultures for the common STDs are definitely recommended and will be outlined in a moment. This is obviously to rule out concomitant infections.

Concerning pregnant women in the trials, we feel that pregnant women may be included after the first trimester unless drug safety is a specific issue.

[Slide.]

Inclusion criteria will be postmenarchal women with a diagnosis of bacterial vaginosis based on all of the

four findings listed here. The characteristic discharge is the thin, homogeneous adherent discharge with minimal or absent inflammation. This may typically be somewhat milky in appearance and is often associated with a fishy odor or a malodor. Vaginal pH greater than 4.5, greater than 20 percent clue cells on the wet mount--and I will go into that further in a moment--and then a positive whiff or amine test with the addition of one drop of 10 percent KOH to the wet mount.

I will make a quick comment here. These are really your classic Amsel criteria from 1983, however, by Amsel's criteria, you only needed three of the four criteria. We are recommending that all four should be present.

There is one additional change here, and that is just 20 percent clue cells. In Amsel's criteria, clue cells simply needed to be present. In 1988, an article by Dave Eschenbach recommended that 20 percent of the epithelial cells be considered clue cells, and this basically made the diagnosis of BV more specific, and so we have elected to recommend the 20 percent clue cell criteria.

[Slide.]

This is a photograph that is fairly representative of a clue cell. There is a squamous epithelial cell here.

This is from a Gram's stain. The most important characteristic here is that the cytoplasmic border of this cell is really totally obliterated by the adherence of the Gardnerella-like organisms which are small, gram-variable bacilli that obscure the cytoplasmic margins here and give it this "shaggy" appearance.

Many epithelial cells will have the bacteria adherent to the surface of the cell, but the sort of true clue cells should have this shaggy border as we see here, and that is considered to be the clue cell.

[Slide.]

For exclusion criteria, it would be women that have other causes of vulvovaginal pathology. They can be either infectious or non-infectious. Patients who have received any antifungal or antimicrobial treatment--here, we could have some discussion about this--we elected 14 days of entry to the study, pregnant women in the first trimester of pregnancy, and pregnant women later in pregnancy should be excluded as noted or if drug safety concerns exist.

[Slide.]

Drug compliance, which is really the same as Joe outlined for vulvovaginal candidiasis. For a single-dose therapy, the patient should receive all drug, three-day therapy the same, but for a five- to seven-day therapy, the

recommendation was that the subject should receive at least the first three consecutive days of drug.

[Slide.]

Evaluation visits. There is a fairly significant change here. Previously, in the studies that were submitted to the FDA, three studies were recommended or carried out for evaluability. We currently are recommending that there be only two visits, that there be a patient diary, and then an interim telephone contact, which I will talk about momentarily, will be in essence a replacement for the middle visit that used to be done.

[Slide.]

So, the evaluation schedule or visit starts with the entry visit, which will have a focused history and physical exam. We would like data about current and past contraceptive data. The past BV history would be especially the 12 months prior to the study, because one major problem with BV is a recurrence of the infection, and then a history about sexual activity because of the close association between, as I mentioned, new sexual partners or multiple sexual partners with this infection. A pap smear, wet mount, 10 percent KOH whiff test would be done at this visit.

Gram stain is obtained, and the most important, it

should be sent to a central reference lab, and this is for the interpretation of the Gram stain by Nugent's criteria, which we will discuss momentarily. STD cultures for GC, Chlamydia, and Trichomonas should be obtained, and the patient diary card with careful instructions will also be handed out.

[Slide.]

We then would follow up with an interim phone call, specifically, day 7 through 10, after day 1 of the study, and the same with VVC. This phone call is not critical for evaluability, however, it should be very valuable to get data in terms of compliance of the patient for both protocol compliance and filling out the diary on a daily basis. It is an excellent time to check on the subject's response to therapy and recording of adverse events.

[Slide.]

The second visit is the test-of-cure visit, and we are recommending between days 21 and 30 of the study. At this time a wet mount and 10 percent KOH whiff test would be performed, a repeat Gram stain is obtained and sent to the same central reference lab.

The patient's diary card and signs and symptoms are evaluated. STD cultures would be required only if they

clinically indicated meaning the investigator suspected there might be a Trichomonas infection or some other vaginal infection.

Also important, though, is the clinical assessment by the investigator at this visit, because that will have a bearing in terms of the outcome of the study.

[Slide.]

Our therapeutic outcome represents a relatively major change in our recommendations. In the past, there have been three categories, namely, clinical cure, clinical improvement, and clinical failure.

We currently are recommending a composite dichotomous outcome of therapeutic cure or therapeutic failure. I will outline this further now.

[Slide.]

This is based on the combination of clinical outcome with either cure or failure and the Nugent scoring system with an outcome of either cure or failure. The Nugent scoring, I will talk further about that in a moment.

[Slide.]

The clinical cure is based on a return to a normal physiologic discharge basically as determined by the investigator, a negative whiff test, saline wet mount negative for clue cells, and a vaginal pH less than 4.7.

Clinical failure would be those subjects who do not meet the above definition of cure or those subjects where the investigator determination is that the patient clinically still has an infection, or any subjects who used other treatment during the study for bacterial vaginosis.

[Slide.]

The Nugent scoring system, I will now outline, is basically based on a total score which is the sum of three weighted scores of the three morphotypes: Lactobacillus, Gardnerella bacteroides morphotype, and Mobiluncus type.

The quantification scale, which is really zero to 4+, is based on the average number of these morphotypes seen in 10 to 20 oil immersion fields. I don't need to bore you with the actual details there, but a typical normal Gram stain slide would show something like 3 or 4+ Lactobacillus, which for that individual morphotype would give a score of 1, and would show a 1+, say, Gardnerella morphotypes, again a score of 1, and maybe no Mobiluncus or 1 to 2+ score of 1, so the cumulative score in the normal Gram stain would be 3 or less, and that is interpreted as normal.

In BV, we see a shift in the morphotypes where we get very few left of Bacillus types, and a marked increase in the Gardnerella and Mobiluncus types. So, a typical slide Gram stain for bacterial vaginosis might be 1+

Lactobacillus type score is 3, 3+ Gardnerella score is 3, and then, say, a 3 or 4+ Mobiluncus score is 2, so cumulative would be 3+3+2 or 8, and a score of 7 to 10 is considered compatible and diagnostic of bacterial vaginosis.

[Slide.]

I have mentioned the total score, and the usual interpretation is that normal is zero to 3, there is an intermediate category and the classic Nugent score of a score of 4 to 6, and BV is 7 or more.

However, despite this intermediate category here, we are recommending that a score greater than 3 can be interpreted or considered abnormal, which would be used for entry criteria, and that a score of zero to 3 would be considered normal and be used as a cure criteria at the end of the study.

[Slide.]

Our overall therapeutic outcome then, as mentioned earlier, depends on both clinical outcome and the Nugent score result. The only combination that we are recommending be considered a therapeutic cure is namely the clinical outcome of cure and a Nugent score of zero to 3.

The next three possibilities--and I won't go over them all--would result in an interpretation of a therapeutic failure, and then the next three would be non-evaluable, and

this deals with the issue that has been brought up many times today, and that is of missing data from the study.

[Slide.]

Evaluability as a cure would include the subjects who have a test-of-cure visit between days 21 and 30, women who uses no other antimicrobial drugs during the study, patients who started the treatment within 48 hours of randomization, no protocol violations, and no other intravaginal products used during Days 1 through 7.

By "intravaginal products," we specifically mean N-9 products, condoms, douches, feminine deodorant sprays, tampons, anything that we feel would be a compounding factor in the interpretation of the data and outcome.

[Slide.]

Evaluability as a failure would be any patients with assessment between Days 4 and 30. Those women who used additional antimicrobial drug at any time during the study, or those subjects where the investigator's clinical determination was that of a failure of treatment.

[Slide.]

There are several options or possibilities for non-evaluability in the study. That would be women with an entry Nugent score of less than 4, those who did not comply with the minimum days of therapy--and it sounds like we need

a little further discussion about that with the advisory committee--women who started the drug greater than 48 hours after randomization, those who used other vaginal products during Days 1 through 7 as just discussed, and those women missing clinical or Gram stain data, which was seen on the table about therapeutic outcome.

[Slide.]

Analytical considerations are that we recommend that analysis on the study results be performed on two populations. Most important in terms of our recommendations would be the per protocol analysis from the strictly evaluable patients. There is also the intent-to-treat group, which is anyone randomized to enter the study.

The primary efficacy variable would be the therapeutic cure at the test-of-cure visit.

[Slide.]

This would be the data from the two multicenter trials. Each trial should be adequately powered to demonstrate the therapeutic equivalence of the test drug to the comparator drug as mentioned for the per protocol evaluable population.

We recommend using the 95 percent confidence interval around the difference in the two therapeutic cure rates.

[Slide.]

Just to summarize briefly on the major changes, we do have a new entry criteria of the Nugent's score that is greater than 3, and not listed here as a bullet would be need for having all four of the clinical findings, namely, the characteristic discharge, pH greater than 4.5, the positive whiff test, and the greater than 20 percent clue cells present.

The number of visits is changed to two instead of three, but we have added the interim phone call as discussed, and we are proposing a new therapeutic outcome based on clinical findings and the Nugent's score, and we are proposing to eliminate any improvement or intermediate category.

[Slide.]

We have three questions for the advisory committee. The first is: Does the advisory committee agree with the proposed use of the Nugent scoring system and the criteria that an overall therapeutic cure is based on a combination of clinical outcome and the Nugent scoring system?

We feel that this is an excellent way to analyze and look at the data because especially with the Nugent scoring system, it is standardized. Comparative means you

potentially can compare the Gram stain slide from the entry visit with the test-of-cure visit for each patient if there is a question of a real true change in the person's microflora, and it will be done in a central lab..

Nugent scoring does allow for an unbiased interpretation. It allows strict criteria for both diagnosis and meeting entry into the study and for cure, and we have proposed that the improvement category be eliminated from the interpretation.

[Slide.]

Our second question is: Does the advisory committee agree with recommendation that our test-of-cure visit occur between Days 21 and 30?

The reason we raise this question is primarily, if you go to the second bullet, the literature is really unclear as to the time it takes for the vaginal flora to normalize to return to normal following antibiotic use.

BV is in essence a derangement of the normal polymicrobial vaginal flora, and with a marked decrease in the Lactobacillus, which is normally the predominant organism in the vagina and with an overgrowth or increase in the Gardnerella bacteroides type and the Mobiluncus type, and often with the Mycoplasma.

So, our question deals with the window for the

test-of-cure visit.

[Slide.]

The third question is: Does the advisory committee believe that placebo-controlled trials are ethical and/or desirable for studies for BV in contrast with the traditional blinded active-controlled studies?

Some reasons to raise this question are that BV, in fact, is a very common infection as noted at the beginning of the talk, and it really does have very mild symptoms, namely, just some increase in discharge and the odor, but it is not associated with a lot of inflammation or a lot of symptoms on the woman's behalf.

If placebo-controlled trials were recommended, there would be the possibility of a better evaluation for adverse events. Certainly, one could measure the actual placebo effect as compared to the inactive control, and we do raise the question that in the future, there may be lower efficacy rates with our new proposed guidelines with the combination of clinical outcome and the Nugent's outcome, so there may be some value in the placebo control if, in fact, in the future, efficacy rates are lowered because of our change in standards and our change in the guidelines.

With that, I conclude my talk and leave it open to the advisory committee and Dr. Craig and Dr. Soper.

DR. CRAIG: Any clarification questions? If not, we will move right on to Dr. Soper.

Committee Presentation

DR. SOPER: The same issues essentially. I reviewed the presentation and would just kind of redefine bacterial vaginosis as Dan has done, and that is, it is a complex alteration of vaginal flora in which Lactobacilli, which normally make hydrogen peroxide and therefore kill all the catalase-negative organisms in the vagina and lower the pH because of lactic acid production, go away for some reason which we do not understand.

What happens when the Lactobacilli go away is that the hydrogen peroxide disappears, the catalase-positive microorganisms overgrow including Gardnerella anaerobic bacteria. They secrete a means which cause a fishy odor, and this phenomenon is defined as bacterial vaginosis.

With respect to study considerations, I just want to make a very brief comment about pregnancy, and I don't want to dwell there, but clearly this disorder has been associated with adverse reproductive tract outcome in both non-pregnant and pregnant women, and there is now data to suggest that therapy during pregnancy can be protective.

Therefore, we need to study this disease and its treatment in pregnancy and in the first trimester, so any

help that we can give industry to support this kind of work, I think would be very important.

In the bigger scheme of things, the literature about antimicrobial therapy in pregnancy in general in this country is terrible, because of liability issues and maybe it is even worthwhile to convene a special interest panel at sometime in the future about this problem to help study antimicrobial therapy at pregnancy.

Anyway, with that said, on to the inclusion criteria. As far as the composite clinical criteria for the diagnosis of BV, homogeneous discharge, whiff test, clue cells, and pH, the homogeneous discharge has a low specificity and sensitivity. You and I cannot agree what that is. That needs to go away somewhere, because nobody uses it, and it does not need to be part of the inclusion criteria.

The issue with inclusion criteria really is based on enrolling patients that can be confirmed to have the disease by the gold standard, which is the Nugent Gram stain. Therefore, you can be as rigorous or as limited as you would like to be with the clinical composite criteria.

My recommendation is that if you want to be really rigorous, and you want to make sure that all your patients will have Nugent criterias greater than 3, that you can

maintain very rigorous criteria greater than 20 percent clue cells, but I don't think you need to be so rigorous because you will use the Gram stain as the gold standard, and I would say that if you use two out of three criteria with the clue cells being necessary, the presence of clue cells has to be one of the two criteria, so in other words, you could have a abnormal pH and clue cells for whiff test and clue cells, but a pH and a whiff test alone would not be good enough, would be my recommendation for composite clinical criteria, and because you used the Gram stain much like in bacterial infection studies as the gold standard, in other words, if the patient doesn't have the pathogen, she is excluded, is that correct or not?

What happens like if you are treating an infection, and the culture is negative, are those patients dropped?

DR. DAVIS: Yes.

DR. CRAIG: That is correct. I mean like for VVC, even is you have some symptoms.

DR. SOPER: So, if your Nugent criteria is, as you point out, is negative, it doesn't make any difference what the composite clinical criteria is, the patient is dropped from the study.

DR. DAVIS: That is right, you are non-evaluable.

DR. SOPER: So, anyway, that would handle inclusion criteria for you, and that would make it easy to enroll patients, and you would still have the gold standard for the enrolled patient being the Gram stain. You can kind of equate culture with Gram stain in this disease. You wouldn't do a culture, but the Gram stain equals the culture if you understand what I am saying.

Extended treatment, test-of-cure, I like. Much like with the vulvovaginal candidiasis issue, the studies of many that have been done, essentially, the evaluation is done immediately following therapy. It gives an inflated view of response to treatment, and I think by delaying the test-of-cure, it gives you a more realistic expectation of what you are really going to be dealing with, and you also need to understand, of course, that BV, because the basic pathophysiology is loss of Lactobacilli and whatever causes this that we have not fixed by treating patients with antimicrobial therapy is that over time, the majority of patients with BV recur, so that if you look at them at six months 50 percent have recurred, at nine months 80 percent have recurred, it is this kind of phenomena.

So, if you then as a corollary suggest that you delay test-of-cure, you are going to get higher failure rates the longer out you go from treatment, so again it's a

little dicey as to what you recommend for your test-of-cure window.

I would recommend again 21 to 45 days, recognizing that if you go 21 to 30 days, you are probably going to have a little bit better therapeutic efficacy because you are going to capture some patients that haven't quite bumped their Nugent score yet, but will bump it by Day 45, et cetera, if you make it even longer.

The therapeutic outcome again is similarly requiring no signs of disease, which is composite clinical criteria and "negative culture," which is really in this case a negative Gram stain, which is very reasonable except you really don't even to need to worry about composite clinical criteria because I don't think you are going to get into a situation where you have composite clinical criteria suggesting BV and a Gram stain being normal. You wouldn't expect to be in that situation.

I mean I guess there may be some data that would support that statement, but to me, in the studies, I like to know that the Gram stain confirms the clinical diagnosis, and if the Gram stain doesn't confirm it, I wonder if maybe the composite clinical criteria wasn't overcalled. Remember my skepticism about the quality of microscopy that is done on vaginal secretions in this country.

In the past, we had an improvement category, as you point out, and you have now recommended that these patients become failures, so that is going to decrease efficacy rates substantially. Is this appropriate?

Well, a Nugent score of zero to 3 is absolutely normal. I mean it is normal, no doubt about it. A Nugent score of 7 and greater is absolutely abnormal, there is no doubt about it. The Nugent score between 4 and 6 clearly is exactly what it says, it is intermediate. Does that mean it is associated with reproductive tract sequelae?

In some cases, yes. It is probably more abnormal than normal, and part of the problem is that bacterial vaginosis is a continuum. So, you have a normal patient with all these Lactobacilli and none of these of these other microorganisms, and you have got the BV patient that has got loads of anaerobes and no Lactobacilli, and then these intermediate patients are somewhere in-between. Where do they really belong, do they belong in the failures, in the abnormal group or not? I don't really know the answer to that except I think they are probably more abnormal than they are normal. I know that is really helpful.

The criteria for cure, the clinical composite criteria I have already alluded to, but again there would be no need to assess the character of discharge because of what

I have already said, it has a poor sensitivity and specificity anyway. Negative whiff is okay, negative clue cells is okay.

A pH of less than 4.7, I would just point out that in your inclusion criteria, you say it needs to be greater than 4.5, so if you are going to include composite clinical criteria, there should be some consistency, so that normal would be less than or equal to 4.5 instead of less than 4.7. I think 4.5 is more reasonable, but herein lies part of the rub, and that is, that many patients that have previously been described as cures never normalize their pH because they can't get that lactic acid production from Lactobacilli that they don't have, although vaginal epithelial cells also supply some of this to try to drop pH. Patients that can resolve their symptoms, and not have clue cells on both composite clinical criteria and Gram stain still will not be able to normalize their pH, so that is going to be a rigorous requirement that may result in some increased failures.

The Nugent criteria of less than or equal to 3 is going to be stringent requirement for cure, because if the patient eliminates or essentially decreases her anaerobes and Mobiluncus, and her Gardnerella, she still may not have any Lactobacilli to get up to 3, so that the best she may be

able to do, despite being symptomatically normal, is a 4 on the Nugent, because she will have no Lactobacilli. She will also have no demonstrable Gardnerella on Gram stain and no Mobiluncus. That still gives her a 4.

So, you need to kind of come to grips with what cure is. Normal is Lactobacilli, but maybe cure is not normal, if you know what I mean. Again, that will increase the overall failure rate and decreased efficacy.

As far as evaluability goes, you point out that the patient should not use intravaginal products for Days 1 through 7, and I would suggest that they probably shouldn't use it throughout the study period, whatever that is, because say, for example, the patient has a response to therapy, douches a week after completion of therapy, and redevelops BV because maybe douching is associated with recurrence of BV or whatever changes vaginal pH may be the instigating element that produces bacterial vaginosis, so I would say nothing in the vagina until the test-of-cure. I guess nothing is not a good statement because obviously, patients are going to continue to have sexual intercourse, but no douching products or other intravaginal products.

In response to your questions, Question No. 1, should we use the Nugent scoring system. I agree that we should use it, and I think it is the gold standard for

confirming the diagnosis, but I think the big question is where do the intermediates belong, are they abnormal and therefore failures, and also the problem that I have alluded to with respect to the Nugent score of 3 maybe being too rigorous, and maybe that needing to be liberalized to a 4.

Question No. 2, what about the recommendation for the extended test-of-cure visit, I like that an awful lot for reasons that I have already stated. I think that in looking at work that has already been done, patients clearly have a relatively prompt resolution of abnormal flora right after therapy.

You point out that we don't know how long it takes patients to normalize flora, and what that suggests is that Lactobacilli overgrow back up to normal concentrations, and therefore we establish the normal vaginal flora, which may never happen in patients that have bacterial vaginosis, and therein lies the rub with late recurrence which can almost be predicted.

No. 3, are placebo-controlled trials ethical, et cetera, I would say placebo-controlled trials are completely ethical in non-pregnant patients, but that those that have failures should be offered standard therapy at the end of the course of their randomized clinical trial, that in pregnancy, that is no longer the case, that the data support

the treatment of bacterial vaginosis in the second trimester.

You could probably design a trial where you randomize patients in the first trimester and then retested them in the second trimester just to make sure that they have cleared their BV in the second. I think that would still be an ethical design.

That is all of my comments.

DR. CRAIG: Dr. Leissa.

Committee Discussion

DR. LEISSA: Dr. Soper, to Question 3, you state that you believe it is ethical for placebo-controlled trials in non-pregnant women. Do you think it is desirable, though, from the perspective of showing efficacy in adequate and well-controlled trials?

DR. SOPER: Yes. I think depending on what falls out with respect to cure definitions, that your efficacy rates may plummet with this disease. I think my suggestion to you is get additional experts besides myself to give some input into this, because I think this requires much more pondering than I have had a chance to do after seeing this.

I would like to know the true efficacy of the new agent, and I think that is best done with placebo-controlled trials.

The trouble with comparative agents now is that all the available treatments given new criteria probably would have lower efficacy rates, probably on the range of 65 percent or something like that. That is a guess, so you would like to see a placebo.

I don't think it would do any harm. As a matter of fact, I think in the long run it does good because you have identified patients with the disease, and you can offer them standard therapy at the completion of the trial.

DR. CRAIG: Why would it drop, do you think it is because of the Nugent criteria?

DR. SOPER: I think that and also early evaluation and overtrials, that if you delay test-of-cure evaluation, that you may see some decrease in efficacy. I don't know if you have the same data from BV studies as you do from VVC studies where you saw early evaluation showed better efficacy than later.

DR. WINFIELD: That is correct. The earlier you evaluate them, the better the results are.

One of the problems we had in terms of the pH, and the reason we accepted a 4.5 pH going in, was because if you have a pH of 4.5 or greater, it was indicative of BV and/or trichomoniasis, but one of the problems we had was in coming out of the trial, the pH was the last thing to go back to

normal, and we found that a lot of patients who had a pH of 4.7 had no symptoms, and basically, the Gram stain was normal, but they were considered as failures because they had not returned their pH to 4.5, so your comment is well taken on the pH and also of putting more emphasis on the Gram stain than on the other clinical findings in terms of inclusion, as well as cures.

DR. DAVIS: The other comment quickly about the intravaginal products, the reason that we felt that it would be definitely a protocol violation if any product was used in the first seven days is that there was really the time of the study drug and clearly the time when you wanted no confounding factors.

The reason why we didn't include it necessarily after that is that in reality and in real life, women are going to have intercourse probably--we don't know--but I mean in Day 10, Day 14, Day 21, may use a N-9 product, a condom, may have their period and use a tampon. I mean that is intravaginal product as far as we are concerned, may use a feminine deodorant spray, who knows, but we really did feel very strongly that no other intravaginal product should be used at Day 1 through 7.

Now, the purest way to do it is no products for the entire study. We would agree with you on that entirely.

It is just a matter of at what point. Again, in the past, many women have been excluded or considered non-evaluable from the clinical trials because they--it is not so much douching--but they did have intercourse on Day 18, and that was a protocol violation, therefore, they were non-evaluable.

So, we are trying to lighten up a little bit on some of those criteria.

DR. SOPER: That is very reasonable, and it reminds me, though, that there should be a caveat about douching within 48 hours of evaluation, because that will alter what you see at the test-of-cure.

DR. CRAIG: I guess I don't have a feel for what percentage of the patients would be cured by one technique, and then what percentage would be picked up with the Nugent criteria that, you know, you didn't pick up with the others by the fact of using them both.

DR. WINFIELD: In the past, most of the studies have not included the Gram stain. They have done the Gram stain, but they have not included them in terms of cures. But once you add the Gram stain to it, it does reduce the number of cures that you see substantially, even though I think this is probably a better marker.

Previously, we have had all kinds of definitions

as to what is a cure. As you can see, you can go in with four different signs and symptoms to include the discharge, the amine odor, the pH, et cetera, and then in coming out they would say that you could have two of the three or three of the four, and it was very confusing as to how you could really appropriately evaluate these patients.

So, we feel that the Gram stain--and I agree with Dr. Soper--is the best way to evaluate BV in terms of making a diagnosis, as well as determining whether or not cure, and second to that I think would be the clue cells.

DR. SOPER: Not only that, but all the epidemiologic work that has been done with BV is really Gram stain based, so if you think about adverse consequences, it is the Gram stain you want to normalize.

DR. CRAIG: But I mean I guess at least we had for the other clinical outcome, we had the pH changes, and if that doesn't come up when the Gram stain normalizes, that could still be called a failure, am I right?

DR. SOPER: Yes, you are right, it would be a failure.

DR. DAVIS: However, on the failure category, the pH, we would allow a pH of 4.7, which is higher than the entry pH of 4.5, because it does take longer for the pH to normalize, so that is why there is a discrepancy there. For

entry, it is greater than 4.5, but for cure, it is greater than 4.7 because of it taking longer for the pH to return back towards normal.

DR. LEISSA: Dr. Soper, you had stated that you felt that the Gram stain was the gold standard, so that for the purposes of entry into the study, if the score was greater than 3, it would be reasonable to include them.

Then, you raised the concern that at the follow-up visit, if someone came back and had not had the Lactobacillus come back, that they could still be considered a failure.

I am just wondering, with that concern in mind, whether we should be dropping the Nugent at the test-of-cure visit, whether Nugent should be used to enter somebody into the study.

DR. SOPER: Well, I think it is going to kind of depend on the severity of the BV. For example, if you have a patient with a score of 4, 5, 6, 7, they have Lactobacilli, then, they are going to be able to score a 3 or less, but if you have somebody that has no Lactobacilli, and that is generally related with much higher scores, you know, the 9, 10's, then, those patients might be more likely to fail given Nugent criteria for resolution.

Again, a lot of this is study design definition,

and as long as it is there across the board, it doesn't really make any difference. You would expect that the agents would work if they are effective similarly, and had the same failure rates based on the severity of the disturbance of the vaginal milieux, so I think that that will all work out given the way that these are designed.

DR. CRAIG: But again, I mean you could go from 10 to 4 and be considered a failure. You could go from 4 to 3 and be considered a cure.

DR. SOPER: I would guess that that would be unlikely, but I am sure it can happen.

DR. ALBRECHT: A corollary I guess to that question is when you have those patients who had follow-up, assuming that we do continue to use Nugent criteria, have 4 to 6 as their values.

Is there any value perhaps having them come back later? Maybe they just take longer to repopulate with the Lactobacilli. Do you think there is value in that or do you think they don't?

DR. SOPER: I don't know the answer to the question, but my overall feeling would be that they might even worsen over time.

DR. CRAIG: Since the clinical trials tend to suggest they get worse the longer you look.

So, how do the rest of the people feel about Question No. 1, about the Nugent criteria?

DR. NORDEN: I am still not clear. Is just one person going to read all the smears, the Nugent? How did you propose to do that?

DR. DAVIS: Our recommendation, I will put it this way. One central lab read them. Now, if you send something to Sharon Hilliard's lab in Pittsburgh, I mean I am not so sure that she sits down and she looks at 600 slides, but it would be her trained microbiologist technicians by sort of standard criteria, and so forth, or if they are sent to Dr. Soper or they are sent to Dave Eschenbach or whatever, but it is one central place that is doing the interpretation, but in fact or in reality, I am pretty sure it is not the same technician or doctor reading all 600 slides.

DR. SOPER: These laboratories have quality control systems in place to make sure that the tech is seeing what the microbiologist is suggesting, so they have a high degree of inter-observer reproducibility and reliability.

DR. DAVIS: And that has been clearly shown in the literature. There are many studies, excellent studies showing the inter- and intra-observer reliability and reproducibility of the Gram stain and Nugent's criteria.

DR. CRAIG: Could you see eventually coming down the line, as you collect more data, and see how the two things behave, that you might end up just picking one of them, I mean specific like the Nugent criteria?

DR. DAVIS: Yes. As Dr. Soper said, the gold standard really is sort of the Nugent's criteria, whereas, the clinical findings, I mean you can get a positive whiff test with Trichomonas, and then you can get it also I think when you don't have BV by other criteria.

You can get pH changes because of douching, because of intercourse, and so forth. Those are somewhat objective, but also somewhat subjective, the interpretation, but the Nugent's criteria with the Gram stain, it is more objective and reproducible.

DR. CRAIG: If it is gold standard, then, by adding in the clinical trials, all we are doing is potentially reducing the value of the gold standard? I mean if we are going to fail some of those, and you need to be success on both, are we sort of reducing the value of the gold standard?

DR. DAVIS: It is a good point.

DR. SOPER: I was just going to say I mean I am not sure you need composite clinical criteria at the end of the study unless it is to identify people that going to be

candidates to be offered traditional therapy and that you could go by the Gram stain as the only test-of-cure.

DR. WINFIELD: I would agree because if you have a normal Gram stain, all the others are going to be normal.

DR. CRAIG: That is what I didn't have a feel for, for how often would you have it where the pH hadn't come back yet, and therefore you were calling them a failure because the Gram stain had normalized, everything else had normalized, but the pH wasn't back up yet, and therefore we are calling that a failure.

DR. WINFIELD: Right. If you get what is defined presently as a normal Nugent criteria, that is, zero to 3, all of the others will be normal, at least 98 percent of the time.

DR. CRAIG: Yes.

DR. ALTAIE: Sousan Altaie, FDA. I wanted to ask Dr. Soper for a clarification as far as if you have a Nugent score of greater than 3, you really don't need 25 percent of the epithelials being a clue cell on the wet mount, and I am under the impression then you are saying that the presence or absence is enough, not necessarily 20 percent being clue cells.

DR. SOPER: If you want to be rigorous with composite clinical criteria, you would put in a proportion

of epithelial cells as being clue cells because you can have an occasional clue cell drift by as you do microscopy, and if it is 1 out of 100,000 cells, it is like that is not BV.

Also, Lactobacilli can adhere to epithelial cells and kind of be a false positive clue cell. So, for a sophisticated microscopist, I don't worry about the ability to make the diagnosis of BV based on clue cells, but not everybody is a sophisticated microscopist, and study nurses do a lot of the work in this area, and therefore, if you are going to use composite clinical criteria, you want a high degree of reproducibility with respect to what you get on Gram stain.

You make some sort of arbitrary cutoff with respect to the proportion of epithelial cells that need to be clue cells. The higher it is, the more reliable your Gram stain is going to reproduce the results of your composite clinical criteria.

If you use the Gram stain as the gold standard, it doesn't make any difference because you may enroll a patient that you think has 2 percent clue cells, and if she doesn't meet Nugent's criteria, she is gone, so it is a moot point.

DR. CRAIG: Yes, Carl.

DR. NORDEN: David, I just need to pursue something with you. The Nugent score fascinates me. I can

understand using it as a diagnostic criteria. It seems pretty clear that if it is greater than 3, that that is fine.

But in terms of cure, I came back to your point earlier that you may not get to 3, and yet be doing well. Bill's point about going from 8 or 7 to 4, I mean should there be, rather than absolute cutoff, an addition to an absolute cutoff as an alternative? Should there be a relative decrease in the Nugent score, does that make sense?

DR. SOPER: I would have to go back and look at Nugent's paper, because I am not exactly sure how they came up with a 3 and then a 4 to 6. Did you guys look at that critically, because that would help with answering that question.

DR. WITTES: But that wouldn't address the Lactobacillus problem. It seems to me there is two different problems. One is the problem that you alluded to, the big job versus the little job, but then there seems to be a threshold over 4 to 3, where there is a qualitative change, and so it seems to me that there is two different problems.

DR. WINFIELD: Previously, what had happened in terms of evaluating these patients, they really weren't determined as cures or failures, but they were determined as

successes or not, and what happened--and this is in his paper--there was clearly a definition of what was cured, and those patients had zero to 3, or clearly those patients who had disease was 7 to 10.

There was the intermediate category from 4 to 6, and what would happen, what he did was he would look at the Gram stain and then go along with the clinical symptoms and determine whether or not they had a success.

So, they didn't say that they were cured or a failure, but they were clinical successes, and this is what we are trying to get away from, is that intermediate category, what would be considered a success, and either classify those patients as either a cure or failure, and this is why we are wrestling with what criteria should we use.

If we use the Gram stain, zero to 3, we would know that, but there will be a lot of them that would probably unjustly classified as failures unless we can add some clinical criteria, as well.

DR. SOPER: As I mentioned before, I think this one issue is dealing with the intermediates, which is kind of the fulcrum of all of this discussion. It is so key that I think you are going to have to convene or talk to a number of people to try to get the best information available and

make a decision.

DR. YEADON: Arnold Yeadon, medical consultant.

From time to time I advise various companies about running clinical trials. Right now I have one company which is involved or about to become involved in a BV trial.

They have been having some difficulty with investigators in the field. If you say, well, we really can't get this 20 percent clue cell, is it really important to have 20 percent. I am somewhat encouraged to hear Dr. Soper say--and I think this is what you said--that if we have Nugent as the gold standard, and if we have "the presence of clue cells," and two of the others, the whiff test and the pH, that we really have adequate information to enter a patient into a clinical study.

Is that really what I am hearing you say?

DR. SOPER: Yes, it is. Again, remember that the 20 percent is more related to the rigor of the composite clinical criteria. If you were just doing a study with composite clinical criteria, I would suggest to you that you would use the 20 percent breakout, but if you are backing that up with a confirmatory Gram stain, I think that is very acceptable.

See, you are really using it in this regard, the composite clinical criteria to accession patients that then

are confirmed to have the disease the Gram stain.

DR. YEADON: That is most encouraging. Thank you.

DR. DAVIS: I would like to make just one quick point there. The clue cells you could see immediately right in the office on a wet smear, whereas, the Gram stain certainly by what we are recommending is it should be sent to an outside reference lab, so you are not going to have the data back from the Gram stain with a Nugent scoring for I mean it might take two, three, four days, or it might just be done a month later.

So, that is the one advantage of saying you would like the clue cells for entry or in randomization because that can be seen literally at the time of the initial study visit.

DR. SOPER: What you would want to do with your center is to make sure that if your investigators are enrolling with a relatively low proportion of clue cells that you are confirming the diagnosis most of the time.

DR. YEADON: Yes, of course, and obviously, we would be willing to accept the proportion of patients who would be non-evaluable because they didn't meet the Nugent criteria.

DR. LEISSA: This question I would ask out of ignorance truly really about the original paper from Nugent,

but it is the Nugent criteria, the Nugent score is the gold standard, is that really more for diagnosis in terms of somebody having a condition and whether we should really say it is also the gold standard for evaluating efficacy? Is that really a fair statement?

DR. SOPER: Again, I guess the way I look at this is that it is much like looking at the eradication of a pathogen for an infection that you diagnose clinically, and it is nice to have, you know, confirmation of the resolution of symptoms with normalization of an exam and a negative culture or, in this case, a negative Gram stain.

DR. CRAIG: Other comments? Any other questions that the FDA has?

DR. DAVIS: Is there any further discussion about the drug compliance, number of days that the drug should be taken by the subject? I know we really did discuss that before with VVC, but does anybody have any additional comments about our recommendations there?

DR. WITTES: I would just to echo my comments, and sort of more generally, if it were possible to redefine the inevaluables in some way, so that there weren't so many of them. I get nervous obviously when it looks like there can be lots of ways of becoming inevaluable, however that can be done to tighten it up would be good.

DR. LEISSA: So, you would you propose--I am just throwing out an idea that if a patient just takes one dose, that they should be considered evaluable--that is what I am asking.

DR. WITTES: I don't know this field, but my gut feeling is as many as you can get in is key, and then back from the beginning of the day, if it turns out that the study, that the compliance is really horrible, then it means that you really haven't been able to evaluate the product. But I think that really then speaks to if you can do a placebo trial, how much better that would be.

DR. MURRAY: It seems like there is a little bit of internal inconsistency, but it is just one of those picky things. It is like evaluable, it started within 48 hours, but it has to be the first three consecutive days or something. I think the way I looked at it, I think those two are mutually incompatible. You couldn't have completed it.

DR. LEISSA: To clarify that, the idea is that when you would be identified for the study, within 48 hours, you would start taking the drug, but then you would have to take it for the next three days, consecutive days of therapy. That is how it is supposed to read.

DR. CRAIG: Dr. Reller.

DR. RELLER: In considering placebo-controlled trials, what was the basis for approval, what kind of trial design, the basis for approval of clindamycin and metronidazole?

DR. WINFIELD: Well, the basis for approval of clindamycin, which was the first vaginal product that was approved, there was no FDA-approved product for the treatment of BV at that time, and therefore it was FDA's decision that the trials that were done, they did four trials, in fact--one was an off-label use of oral metronidazole, and the other one was the use of Sultrin cream, which has an indication for the treatment of Gardnerella vaginalis, and the other was placebo.

So, we looked at the placebo trial and the Sultrin trial, and the decision for approval was made on those two trials.

DR. RELLER: And efficacy was based on clinical endpoints of some kind, I mean no Gram stain, no Nugent, et cetera?

DR. WINFIELD: It was using the four criteria that were mentioned about the discharge, the whiff test, the pH, and the clue cells. Gram stain at that time had not come in as the gold standard.

DR. RELLER: And the clue cells were presence or

absence on wet mount?

DR. WINFIELD: That is correct.

DR. RELLER: It seems to me--and I particularly like David's recommendation--I mean if there are approved things now, why not have those as a comparator? When one looks back at the history of how they were approved, it seems to me that a placebo-controlled trial with the follow-up for symptomatic patients of something that is already available would be an important thing to do.

DR. WITTES: Can I actually go back two tabs? I mean some of these things, an argument it seems to me could be used in the two previous, the vulvovaginal candidiasis, and even otitis media. I mean what I was hearing was two 75 percents, that 75 percent of women get VVC sometime in their life, which suggests that it is not really so horrible, so if there were a limit to mild and moderate cases in a placebo controlled, that might be feasible, and I think I heard that 75 percent of otitis media cases spontaneously get better. Did I hear that?

DR. DAVIS: Yes.

DR. WITTES: But again, the question is--I would throw it on the table.

DR. DAVIS: I think one simple difference with the vulvovaginal candidiasis is the women are really

symptomatic, whereas, in BV, as I mentioned, the only symptoms are the odor and an increased discharge, but we are not talking about a crazy amount of itching, burning, I mean just much more of a physical discomfort, if you will, so that to do a placebo-controlled trial for candidiasis, I think you would have a lot of unhappy customers there.

DR. WITTES: Well, no, I meant limited to those that were in the pretty mild range.

DR. CRAIG: Any other questions?

DR. ALTAIE: I am kind of confused from the answers that I got from Dr. Soper when I said 25 percent is not necessary, and then the gentleman behind me got the answer that you don't need 25 percent of epithelials being clue cells.

I understand that they are part of our inclusion criteria at this point. If you don't have 25 percent clue cells on your wet prep, you are not to be enrolled. Now, I realize that we can't--

DR. SOPER: I would drop that criteria if you are going to use Nugent.

DR. ALTAIE: That is what I wanted to clarify.

DR. SOPER: If you are going to confirm your entry criteria with Nugent--

DR. ALTAIE: All right. We need to drop that.

DR. RELLER: I am embarrassed to say, I mean I know that when we get specimens, the Gram stain is done on everybody, and it is compared with Sharon Hilliard's published picture, I mean saying this is a positive, this is a negative, to it's visual yes/no.

People are trained. I mean there is a rigorous reading. You are certified to read these things is what I am trying to say, but they are read against a standard.

DR. SOPER: You literally count like organisms.

DR. RELLER: What is the relationship between the 20 percent or simple presence or absence in the Nugent, because if one is using the Nugent ultimately as the basis for evaluation, mean that the patients--you have a point of randomization, you need something to be able to randomize the patients, but later they are going to be excluded if the Nugent score were greater than 3, so it is a difference in Nugent score that is going to be the ultimate arbiter for assessment of efficacy.

If one does not have Nugent straight away, then, the criteria for entry, if they are too loose, you waste a lot of money and effort. If they too tight, you exclude patients that were perfectly evaluable in terms of assessment of the drug.

So, what is the relationship between the 20

percent or presence or absence--what do you suggest is a reasonable cluster of clinical criteria on which to randomize the patients that ultimately are retained in the trial based on their entry Nugent score?

DR. SOPER: My recommendation would be two out of the three clinical criteria, the presence of any clue cells, a pH that is abnormal, greater than 4.5, and a positive whiff test, but the clue cells have to be present.

If you do work in this area, what you recognize relatively promptly is what a clue cell is, and you generally won't make the mistake of calling a false positive clue cell or just because of your excitement about the potential enrollment, you know, say oh, there goes one clue cell, because if you have at least one additional criteria and a clue cell, that should correlate with an abnormal Nugent, although it only may be in the intermediate range, I would say greater than 90 percent of the time.

DR. RELLER: A simpler way of putting it, the entry criteria would include the presence of clue cells with one or more of these other two criteria.

DR. SOPER: And the reason for not including the homogeneous discharge is because I don't know exactly what that means, and we can't communicate what that means.

DR. RELLER: It would sharpen it up, and it is in

concert with clinical practice. I mean people are looking for clue cells. The laboratory is issuing a result, it is positive or it is negative as a surrogate for saying this patient has or doesn't have laboratory-corroborated or confirmed BV.

DR. CRAIG: Hopefully, what the laboratory is doing is giving them the Nugent score, and not just a clue cell, but you are right.

So, is everyone satisfied? I guess we will close the day's session, and see people again tomorrow.

Thank you.

[Whereupon, at 5:40 p.m., the proceedings were recessed to be resumed at 8:00 a.m., Thursday, July 30, 1998.]